

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:39:00 ; Search time 7452 Seconds

(without alignments)
16919.604 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909
Sequence: 1 ggaagaagtttgtatgcctc.....cagcgaatcggcgcgtgcag 2909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sbs:*
26: em_un:*
27: em_vl:*
28: em_vr:*
29: em_vr:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_nam:*
37: em_hcg_vrt:*
38: em_hcg_hum:*
39: em_hcg_mus:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2909	100.0	2909	1	VCH6514	AJ006514 Vibrio ch
2	2909	100.0	2909	6	AX009481	AX009481 Sequence
3	2870.6	98.7	9931	1	AE004153	AE004153 Vibrio ch
4	1390.8	47.8	1410	1	VCH7146	VY17135 Vibrio chol
5	1222	42.0	1222	6	AX009483	AX009483 Sequence
6	1216.8	41.8	1236	1	AY143429	AY143429 Vibrio ch
7	1187	40.8	1591	1	VCA110968	AJ010968 Vibrio ch
8	1052.4	35.2	298900	1	AP005074	AP005074 Vibrio pa
9	1035	35.6	301442	1	AE016798	AE016798 Vibrio pa
10	1028.6	35.4	249150	1	AP005332	AP005332 Vibrio vu
11	841.4	28.9	852	6	AX020429	AX020429 Sequence
12	838	28.8	838	6	AX009482	AX009482 Sequence
13	814.8	28.0	10719	1	AE006043	AE006043 Pasteurel
14	775.2	25.6	11545	1	AP064791	AF064791 Pasteurel
15	744.2	25.6	11545	1	U32772	U32772 Haemophilus
16	744.2	25.6	110000	6	AR274513_09	AR274513 Haemophil
17	433.8	14.1	873	6	AE017153	AE017153 Haemophil
18	410.8	14.1	873	6	AR377276	AR377276 Sequence
19	352.2	12.1	304250	1	AX294152	AX294152 Pirellula
20	329.4	11.3	10687	1	AE015577	AE015577 Shewanell
21	314.8	10.8	304558	1	AE017151	AE017151 Haemophil
22	292.2	10.0	807	6	AR139598	AR139598 Sequence
23	280.8	9.7	10474	1	AE004472	AE004472 Pseudomon
24	280.6	9.6	10029	1	AE013918	AE013918 Yersinia
25	280.6	9.6	208050	1	AJ414144	AJ414144 Yersinia
26	278.2	9.6	342905	1	BX571861	BX571861 Phototrab
27	278.2	9.6	349980	6	AX770908	AX770908 Sequence
28	273	9.4	10242	1	AE005511	AE005511 Escherich
29	273	9.4	266558	1	AP002563	AP002563 Escherich
30	265.8	9.1	891	6	AR377359	AR377359 Sequence
31	264.2	9.1	305325	1	AE016765	AE016765 Escherich
32	261.4	9.0	310613	1	AE016875	AE016875 Pseudomon
33	260.8	9.0	10405	1	AE000366	AE000366 Escherich
34	260.2	8.9	12943	1	AE015298	AE015298 Shigella
35	260.2	8.9	292906	1	AE016988	AE016988 Shigella
36	259.6	8.9	71128	1	ECU29581	U29581 Escherichia
37	258	8.9	1473	1	STY1GTX	L13259 Salmonella
38	258	8.9	21583	1	AE008838	AE008838 Salmonella
39	256.8	8.8	879	6	AR386236	AR386236 Sequence
40	256.4	8.8	230050	1	AL627277	AL627277 Salmonella
41	256.4	8.8	301311	1	AE016843	AE016843 Salmonella
42	248.2	8.5	876	6	AX188981	AX188981 Sequence
43	248.2	8.5	1310	1	ECOLGT	U12289 Escherichia
44	240.8	8.3	301086	1	AE016964	AE016964 Coxidia
45	232.6	8.0	303438	1	AE017212	AE017212 Geobacter

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
VCH6514	VCH6514	Vibrio cholerae 1gt and thya genes.	AJ006514	1gt gene; prolipoprotein diacylglycerol transferase; thya; thymidylate synthetase.	Vibrio cholerae	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.	Carlén, N.I.A., Nilsson, A., Todorovic, M., Holmgren, J. and Lebens, M.	Characterisation of the 1gt/thya locus from Vibrio cholerae	

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2909)
AUTHORS Carlin,N.I.A.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1998) Carlin N.I.A., Department of Molecular
Biology, SBL Vaccin AB, PMB, JV 30, Stockholm, S-105 21, SWEDEN
Location/Qualifiers
FEATURES
source
1..2909
/organism="Vibrio cholerae"
/mol_type="genomic DNA"
/strain="US1569"
/db_xref="taxon:666"
14..829
/gene="lgt"
14..829
/gene="lgt"
/codon_start=1
/transl_table=11
/product="prolipoprotein diacylglycerol transferase"
/protein_id="CAA07072.1"
/db_xref="GI:3201564"
/db_xref="GOA:O87077"
/db_xref="SPTREMBL:O87077"
/translation="MPQYLQFPNIDPVLFSIGPLAVRWYGLMVLVGLFAMWLANR
ADRAGSGWTRQVSDILFAGFLGVVIGGVVIFYNEDLFLADPLYLEKWTGMSF
HGLGLVITAMFWYARKNQTFFGVADFWAPLVPPLPGLGMRIGNFMNSLMGRVDP
WAFVPPNGEPLFRHPSQLYEFALFEGVLEFLLNWFIGKRPGLGSVSGFLAGYGFRE
LVEVYRDPNQLGFLGFLSMGQILSLPWWIIGILMVMWSYRGLYQDRVAK"
669..775
/gene="thya"
669..775
/gene="thya"
691..696
/gene="thya"
/join(828..832,839..1690,1725..1765)
-10_signal
/gene="thya"
828..832
/gene="thya"
839..1690
/gene="thya"
/codon_start=1
/evidence="experimental"
/transl_table=11
/product="thymidylate synthetase"
/protein_id="CAA07073.1"
/db_xref="GI:3201565"
/db_xref="GOA:O86108"
/db_xref="SWISS-PROT:O66108"
/translation="MKQYLDLCORIVDQGVWVENERTGKRLTVINADLTVDVGNNOF
PLVTRKSPFMAVAELIGYIRGYDMAADPRLQKTMNDANLNQAMLNPRKGED
DMGRVYGVGRAMAKPDGHIQDLKKIVDLSRGVDREILNPNPGEFHHGCRPC
MYSHHSILGDTLYLNSLQKSCDVLGALNFMVQYVPLALMAQITGKKRGLAKYKIV
NHHYQDQLMRDVQLRKPEPPAPQFHNPKIKTLDLETVTDLDDPDVTSYQFHPD
IQYPSV"
1082..1084
/gene="thya"
/note="In trimethoprim resistant mutant US1569 4.4, stop
codon"
/replace="tag"
1725..1765
/gene="thya"
ORIGIN
Query Match 100.0%; Score 2909; DB 1; Length 2909;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGAGCTTTGATAGCTCAGGCTATCGAGTTTCCAAATATTGACCCGATATTGT 60
DB 1 GAGAAAGTTTGTATGCTCAGGGTATCTGCAATTCCCAATATTGACCCGATATTGT 60
QY 61 TTGCATCGGCCCTCTAGCGGTGCGCTGATGAGCTTGATATTGGTGGGTTCCCTTT 120

DB 61 TTGCATCGGCCCTCTAGCGGTGCGCTGATGAGCTTGATATTGGTGGGTTCCCTTT 120
QY 121 TGCATATGTTGGCCCAATCCGCCGAGCGGATCGCGCGGAGTGGTGGACGCGTAGCA 180
DB 121 TGCATATGTTGGCCCAATCCGCCGAGCGGATCGCGCGGAGTGGTGGACGCGTAGCA 180
QY 181 AGTCTCTAGCTTTGATATTCGCGGCTTTTAAAGTGTAGTATTCGGTGGCCGAGTTGGTTA 240
DB 181 AGTCTCTAGCTTTGATATTCGCGGCTTTTAAAGTGTAGTATTCGGTGGCCGAGTTGGTTA 240
QY 241 TGTGATCTTTCAACAATTTGATCTGTTCCTTCCTGACCCCTCTTATTTAAAGTGTG 300
DB 241 TGTGATCTTTCAACAATTTGATCTGTTCCTTCCTGACCCCTCTTATTTAAAGTGTG 300
QY 301 GACTGGCGGCATATGCTTCCACGCGGCTTATTTGGTGTGATCAACGCGCATGTTCTGTGA 360
DB 301 GACTGGCGGCATATGCTTCCACGCGGCTTATTTGGTGTGATCAACGCGCATGTTCTGTGA 360
QY 361 TGCGGGTAAAAACCAACGACACCTTCTTTGTTGGTGGCGGATTTGTTGCCCTTTAGTCC 420
DB 361 TGCGGGTAAAAACCAACGACACCTTCTTTGTTGGTGGCGGATTTGTTGCCCTTTAGTCC 420
QY 421 ATTGGTGGGGAATGGAGCGATTCGGTAACCTTAATGAATAGTGAATTTGGGGAGCGAGT 480
DB 421 ATTGGTGGGGAATGGAGCGATTCGGTAACCTTAATGAATAGTGAATTTGGGGAGCGAGT 480
QY 481 AACGGATGTCCTTGGGCTTTTGTATTCCTTAATGTTGGCCCACTGCGCGGCATCTTTC 540
DB 481 AACGGATGTCCTTGGGCTTTTGTATTCCTTAATGTTGGCCCACTGCGCGGCATCTTTC 540
QY 541 ACAGCTTTATGAATTCGCTTTAGAAAGCGGTGCTGTCTTTATTTCTTAATTTGTTTAT 600
DB 541 ACAGCTTTATGAATTCGCTTTAGAAAGCGGTGCTGTCTTTATTTCTTAATTTGTTTAT 600
QY 601 TGTGTAACCTGTCGCTGAGGAGCGATTCGGATCGGATGTTTAAAGTGAATCGGTAATT 660
DB 601 TGTGTAACCTGTCGCTGAGGAGCGATTCGGATCGGATGTTTAAAGTGAATCGGTAATT 660
QY 661 CCGCTTCTTGGGAATAGTCCGTGAGCCAGATGCTCAGTTGGGCTGTTTGGTGGCTT 720
DB 661 CCGCTTCTTGGGAATAGTCCGTGAGCCAGATGCTCAGTTGGGCTGTTTGGTGGCTT 720
QY 721 CATTTCAAATGGGCAAAATCTCTCTTAACTTAATGATGATCAATCGGATTTTGAATGATG 780
DB 721 CATTTCAAATGGGCAAAATCTCTCTTAACTTAATGATGATCAATCGGATTTTGAATGATG 780
QY 781 TTGGTCTTAAACAAGCCGGTGTATCAACAAGCCGTGTACACCAAAATAGGGTATAGT 840
DB 781 TTGGTCTTAAACAAGCCGGTGTATCAACAAGCCGTGTACACCAAAATAGGGTATAGT 840
QY 841 GAAACAGTATTAGATCTTTGTCAAGCGCATCGTCAAGAGTGTGGTGAAGTGA 900
DB 841 GAAACAGTATTAGATCTTTGTCAAGCGCATCGTCAAGAGTGTGGTGAAGTGA 900
QY 901 ACGAAACGCGCAAGCGTGTGTGACTGTGATTAATGCGGATTTGACCTTACGATGTCGCA 960
DB 901 ACGAAACGCGCAAGCGTGTGTGACTGTGATTAATGCGGATTTGACCTTACGATGTCGCA 960
QY 961 CAATCAGTTTCTCTAGTGAATCAACAAGAGTTTGGAAAGTGTGCCGTGACCGGAGT 1020
DB 961 CAATCAGTTTCTCTAGTGAATCAACAAGAGTTTGGAAAGTGTGCCGTGACCGGAGT 1020
QY 1021 GCTGGGCTAATTCGTGGTAAAGATTAATGCGGAGTTTGGCCATTAAGTACCAAAAC 1080
DB 1021 GCTGGGCTAATTCGTGGTAAAGATTAATGCGGAGTTTGGCCATTAAGTACCAAAAC 1080
QY 1081 CTGGGATCTAATGCAATTTAAACAAGCATGCTCAACAATCTTTACCGTAAGAGTGA 1140
DB 1081 CTGGGATCTAATGCAATTTAAACAAGCATGCTCAACAATCTTTACCGTAAGAGTGA 1140
QY 1141 GGATGACATGGGACCGGTGTATGTGTTCAGGGTGAAGCTTGGGCTAAGCCTGATGTGG 1200
DB 1141 GGATGACATGGGACCGGTGTATGTGTTCAGGGTGAAGCTTGGGCTAAGCCTGATGTGG 1200

QY 1201 TCATATGACGAGTTGAAAAAGATTGTTGATGATTTGAGCCGTGGCGTTGATGACCGAGG 1260
 DB 1201 TCATATGACGAGTTGAAAAAGATTGTTGATGATTTGAGCCGTGGCGTTGATGACCGAGG 1260
 QY 1261 TGAATTTCTTAATCTTACACATCCGGGTGAATTTTCATGATGGGGTGTGGCCCTTGCAAT 1320
 DB 1261 TGAATTTCTTAATCTTACACATCCGGGTGAATTTTCATGATGGGGTGTGGCCCTTGCAAT 1320
 QY 1321 GTACAGCATCATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTCAGGCTTC 1380
 DB 1321 GTACAGCATCATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTCAGGCTTC 1380
 QY 1381 ATGTGATGTGCTTGGGGTGAATTTCAACATGGTGCAGGTTTATGTTCCTTGCGCT 1440
 DB 1381 ATGTGATGTGCTTGGGGTGAATTTCAACATGGTGCAGGTTTATGTTCCTTGCGCT 1440
 QY 1441 GATGGCAAGATCAACAGGAAAAAGCCGGCTGGCGTATCAACAGATGCTCATGCGCA 1500
 DB 1441 GATGGCAAGATCAACAGGAAAAAGCCGGCTGGCGTATCAACAGATGCTCATGCGCA 1500
 QY 1501 CATTTACCAAGATCACTGAAATTGATGGCGGATGTCAGCTTAAACGTGAGCCATTCCTC 1560
 DB 1501 CATTTACCAAGATCACTGAAATTGATGGCGGATGTCAGCTTAAACGTGAGCCATTCCTC 1560
 QY 1561 AGCGCTCAGTTCCATATCATCTCAAGATTAAACACTGACAGATTTGGAACTTGGGT 1620
 DB 1561 AGCGCTCAGTTCCATATCATCTCAAGATTAAACACTGACAGATTTGGAACTTGGGT 1620
 QY 1621 CACTTTGATGATTTTGAAGCTCACCGGATATCACTTCCAGATCTTATTCATATCCGCTT 1680
 DB 1621 CACTTTGATGATTTTGAAGCTCACCGGATATCACTTCCAGATCTTATTCATATCCGCTT 1680
 QY 1681 TTCAGCTAATCCGATATCAGCGGATGCTTGAATGGTTTATATATATATATATATATATAT 1740
 DB 1681 TTCAGCTAATCCGATATCAGCGGATGCTTGAATGGTTTATATATATATATATATATATAT 1740
 QY 1741 CGAAGTTCGGGAGCTTTTATATACAGATGATCTTTAAAGCTTAAAGCGTTAGGGCAAG 1800
 DB 1741 CGAAGTTCGGGAGCTTTTATATACAGATGATCTTTAAAGCTTAAAGCGTTAGGGCAAG 1800
 QY 1801 AATGCTGCGGGGATGACGACAAACACACCAATATGATCTACACCAACCATTTTGGCT 1860
 DB 1801 AATGCTGCGGGGATGACGACAAACACACCAATATGATCTACACCAACCATTTTGGCT 1860
 QY 1861 CTTTACAGCCCAAGTTGATGATGCTCAGACCTTTTAAATGAGCATGTTCCGTTAAGAAAG 1920
 DB 1861 CTTTACAGCCCAAGTTGATGATGCTCAGACCTTTTAAATGAGCATGTTCCGTTAAGAAAG 1920
 QY 1921 AATACGTTAATCAAGACCGTACCATCAAGTTAAAGCTTAAAGTGCACCGCATTTTG 1980
 DB 1921 AATACGTTAATCAAGACCGTACCATCAAGTTAAAGCTTAAAGTGCACCGCATTTTG 1980
 QY 1981 CAGAGCAAAACACGCAAACTCACACAGACAGCGGTTGGCGGACAGACAGCTAATATCA 2040
 DB 1981 CAGAGCAAAACACGCAAACTCACACAGACAGCGGTTGGCGGACAGACAGCTAATATCA 2040
 QY 2041 AGTGCCTAATGTTGCACTTAAGTTAATGGTTAATGATTTTACGCACTTTTACGACGCGCA 2100
 DB 2041 AGTGCCTAATGTTGCACTTAAGTTAATGGTTAATGATTTTACGCACTTTTACGACGCGCA 2100
 QY 2101 GCCCAAGAGAGAAACATTAAGCTGTTGTCGATGTAAGATTAAGTAACTAATCCGTTAAC 2160
 DB 2101 GCCCAAGAGAGAAACATTAAGCTGTTGTCGATGTAAGATTAAGTAACTAATCCGTTAAC 2160
 QY 2161 CACTGTACTGAAAGCAATACCGTGTAGTGGGCTTCGCGCAATGCAATTTTGTGAATTTTC 2220
 DB 2161 CACTGTACTGAAAGCAATACCGTGTAGTGGGCTTCGCGCAATGCAATTTTGTGAATTTTC 2220
 QY 2221 AGGTGGGGGCAACATCAAACTTTTCACTGATTTGGCCATATCCGTTAATGGCGACGAA 2280
 DB 2221 AGGTGGGGGCAACATCAAACTTTTCACTGATTTGGCCATATCCGTTAATGGCGACGAA 2280

QY 2281 AATGTGCAATATCCCAATATGATGTCGACACCAACCGAAAGTATTAACCAATATCCGA 2340
 DB 2281 AATGTGCAATATCCCAATATGATGTCGACACCAACCGAAAGTATTAACCAATATCCGA 2340
 QY 2341 AAGCTGGGTTTAAAGCCCTGATATGACAGTTTGGTAAATGGTTGATTAATCAAAAC 2400
 DB 2341 AAGCTGGGTTTAAAGCCCTGATATGACAGTTTGGTAAATGGTTGATTAATCAAAAC 2400
 QY 2401 TTTTCAATGCTATATGCGCAGTGCAGAGGCGAAAGCAAGCCAGTGTAGACTTTCTC 2460
 DB 2401 TTTTCAATGCTATATGCGCAGTGCAGAGGCGAAAGCAAGCCAGTGTAGACTTTCTC 2460
 QY 2461 TAAATGCAAAATCATATTTCTAGAGTTGAAAGTACACACCGCGAAGATTGAAAA 2520
 DB 2461 TAAATGCAAAATCATATTTCTAGAGTTGAAAGTACACACCGCGAAGATTGAAAA 2520
 QY 2521 ATCTGTGATGTGTCATCTGGCGAAAGCAAGGGAATCTTTCTTATACCGCATATGGCC 2580
 DB 2521 ATCTGTGATGTGTCATCTGGCGAAAGCAAGGGAATCTTTCTTATACCGCATATGGCC 2580
 QY 2581 AAGGCTGACGAGATATGTTGTCAGATGATACCAATATGGCACCCATACCATAGAGAT 2640
 DB 2581 AAGGCTGACGAGATATGTTGTCAGATGATACCAATATGGCACCCATACCATAGAGAT 2640
 QY 2641 CGCGGTTTCAACCGGTTAACCAACCGGCAACGAGACCAATATATGAAAGTCAACGCTGCT 2700
 DB 2641 CGCGGTTTCAACCGGTTAACCAACCGGCAACGAGACCAATATATGAAAGTCAACGCTGCT 2700
 QY 2701 TGAAGATTGAATCAGTCCGTTGCCATTAACCAATCATCAATCTGCAATTGGTGGGA 2760
 DB 2701 TGAAGATTGAATCAGTCCGTTGCCATTAACCAATCATCAATCTGCAATTGGTGGGA 2760
 QY 2761 AGCAATTTGAATAGACTTGGCTTGAATGCGCGGTTGCCATTTAAACGCGTGCAGC 2820
 DB 2761 AGCAATTTGAATAGACTTGGCTTGAATGCGCGGTTGCCATTTAAACGCGTGCAGC 2820
 QY 2821 CATCGGACTGCAAGAGATGTAATATACAGATGAAAGCCAAAGTTTGCCTCAAGTAGGCC 2880
 DB 2821 CATCGGACTGCAAGAGATGTAATATACAGATGAAAGCCAAAGTTTGCCTCAAGTAGGCC 2880
 QY 2881 TTTTGTGTGACGCAATGCGCGCTGCGAG 2909
 DB 2881 TTTTGTGTGACGCAATGCGCGCTGCGAG 2909

RESULT 2
 AX009481 2909 bp DNA linear PAT 06-SHP-2000
 LOCUS AX009481
 DEFINITION Sequence 1 from Patent WO9961634.
 ACCESSION AX009481
 VERSION AX009481.1 GI:9996766
 KEYWORDS
 SOURCE
 ORGANISM
 Vibrio cholerae
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 Vibrionaceae; Vibrio.
 REFERENCE
 1 Carlin, N. and Lebens, M.R.
 Method of producing thy a<->strains of vibrio cholerae, such
 strains and their use
 Patent: WO 9961634-A 1 02-DEC-1999;
 JOURNAL CARLIN NILS (SE); SBL VACCIN AB. (SE); LEBBENS MICHAEL R (SE)
 FEATURES
 location/Qualifiers
 1..2909
 /organism="Vibrio cholerae"
 /mol_type="unasigned DNA"
 /db_xref="taxon:666"
 ORIGIN
 Query Match 100.0%; Score 2909; DB 6; Length 2909;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAGAAAGTTTGTATTCGCTCAGGGGTATTCAGTATTCACAAATATACACCCGATTTGT	60
Db	1	GAGAAAGTTTGTATTCGCTCAGGGGTATTCAGTATTCACAAATATACACCCGATTTGT	60
QY	61	TTGCATCGGCCCTCTAGCGGTGCCTGGTATAGGCTTGATTTGGTGGGTTTCCTTTT	120
Db	61	TTGCATCGGCCCTCTAGCGGTGCCTGGTATAGGCTTGATTTGGTGGGTTTCCTTTT	120
QY	121	TGCTATGTGTGGCCAAATGCGCCAGCGGATGCGCGGCGAGTGAGCTGTGAACGCGTGAACA	180
Db	121	TGCTATGTGTGGCCAAATGCGCCAGCGGATGCGCGGCGAGTGAGCTGTGAACGCGTGAACA	180
QY	181	AGTCCTGACCTTTGATATTCGCGGCTTTTATAGTGATGTATCGGTGGCGAGTTGGTGA	240
Db	181	AGTCCTGACCTTTGATATTCGCGGCTTTTATAGTGATGTATCGGTGGCGAGTTGGTGA	240
QY	241	TGTGATCTTACAAATTTTGATCTGTCTTGCTGCTGACCCGCTTTATTTATCAAGTGTG	300
Db	241	TGTGATCTTACAAATTTTGATCTGTCTTGCTGCTGACCCGCTTTATTTATCAAGTGTG	300
QY	301	GACTGGCGGCAATGTCCTTTCACACGCGGCTTATTTGGGTGTATCACCGCATGTTCTGGTA	360
Db	301	GACTGGCGGCAATGTCCTTTCACACGCGGCTTATTTGGGTGTATCACCGCATGTTCTGGTA	360
QY	361	TGCGCGTAAAAACCAACGACACTTCTTTGTGTGTGGCCGCAATTTGTGTGCCCTTTAGTGCC	420
Db	361	TGCGCGTAAAAACCAACGACACTTCTTTGTGTGTGGCCGCAATTTGTGTGCCCTTTAGTGCC	420
QY	421	ATTCGCTTTGGGGATGGAGCGTATTCGTATACCTTATGAAATAGTGAACCTTTGGGAGCAGT	480
Db	421	ATTCGCTTTGGGGATGGAGCGTATTCGTATACCTTATGAAATAGTGAACCTTTGGGAGCAGT	480
QY	481	AACGGATGTGCTTTGGGCTTTTGTATTCCTAATGTGTGGCCACTGCGCGCCATCTCTC	540
Db	481	AACGGATGTGCTTTGGGCTTTTGTATTCCTAATGTGTGGCCACTGCGCGCCATCTCTC	540
QY	541	ACAGCTTATGAATTCGCTTATGAAAGGCGTGTCTGTCTTATCTTATATTTGGTTAT	600
Db	541	ACAGCTTATGAATTCGCTTATGAAAGGCGTGTGTCTTATCTTATATTTGGTTAT	600
QY	601	TGCTAAACCTGTCCGCTAGGAGCGATATCCGGACTGTTTTAGCTGGAATACGGATCAAT	660
Db	601	TGCTAAACCTGTCCGCTAGGAGCGATATCCGGACTGTTTTAGCTGGAATACGGATCAAT	660
QY	661	CCGCTTCCCTTGTGAATACTGTCGCTGACCAAGATGCTAGTGGGCTGTTTGGTGCTT	720
Db	661	CCGCTTCCCTTGTGAATACTGTCGCTGACCAAGATGCTAGTGGGCTGTTTGGTGCTT	720
QY	721	CATTTCAAATGGGGCAATCTCTCTCTTACCTATGTGTATCATCGGTATTTGATGATGT	780
Db	721	CATTTCAAATGGGGCAATCTCTCTCTTACCTATGTGTATCATCGGTATTTGATGATGT	780
QY	781	TTGGCTTACAAAGCGCGGTTGTATCAAGACCGGTATACACAAATATGGGTAGTTAGGT	840
Db	781	TTGGCTTACAAAGCGCGGTTGTATCAAGACCGGTATACACAAATATGGGTAGTTAGGT	840
QY	841	GAACAGATTTTGAATCTTTTGTACAGCGCATCTGCATCAAGGTGTTTGGTGTGAAAATGA	900
Db	841	GAACAGATTTTGAATCTTTTGTACAGCGCATCTGCATCAAGGTGTTTGGTGTGAAAATGA	900
QY	901	ACGAACGGGCAAGCGTGTGTGACTGTATTAATGCGATTGACCTACAGATGTGGCAA	960
Db	901	ACGAACGGGCAAGCGTGTGTGACTGTATTAATGCGATTGACCTACAGATGTGGCAA	960
QY	961	CAATATAGTTTCTCTAGTGAATCAACGAAAGTITTTGAAAGGTGCGGTAGCGCAATT	1020
Db	961	CAATATAGTTTCTCTAGTGAATCAACGAAAGTITTTGAAAGGTGCGGTAGCGCAATT	1020
QY	1021	GCTGGCTATATTCGTGGTTACGAATAATGCGCGGATTTTGCCCAATATAGGTACCAAAAC	1080
Db	1021	GCTGGCTATATTCGTGGTTACGAATAATGCGCGGATTTTGCCCAATATAGGTACCAAAAC	1080
QY	1081	CTGGGATCTATGTGCCAATTTAAACCAAGCATGTGCTCAACATCTCTTACCGTAAGGTGA	1140

Db	1081	CTGGGATGCTAATCCCAATTTAAACCAAGATGGCTCAACAACTCTTAACGTAAAGGTGA	1140
QY	1141	GGATGACATGGGACCGCGTATGSGTTCAGGGTAGAGCTTGGGCTAAGCTGATGGTGG	1200
Db	1141	GGATGACATGGGACCGCGTATGSGTTCAGGGTAGAGCTTGGGCTAAGCTGATGGTGG	1200
QY	1201	TCATATTGACACGTTGAAAAAGATTGTTGATGATTTAGCGCGTGGCGGTGATGACGAGG	1260
Db	1201	TCATATTGACACGTTGAAAAAGATTGTTGATGATTTAGCGCGTGGCGGTGATGACGAGG	1260
QY	1261	TGAATTTCTTAACCTTCACAAATCCGGGTGAAATTTACATGGGAGTTTGCGCCCTTGAT	1320
Db	1261	TGAATTTCTTAACCTTCACAAATCCGGGTGAAATTTACATGGGAGTTTGCGCCCTTGAT	1320
QY	1321	GTACAGCCATCATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTCAGCGTTG	1380
Db	1321	GTACAGCCATCATTTTTCATTTGCTGGGGGATACCTTGATCTCAACAGTACTCAGCGTTG	1380
QY	1381	ATGTAATGTGCGCTTGGGGTTGAAATTTTCAACATGATGACAGTTTATGTTCCCTTGCGCT	1440
Db	1381	ATGTAATGTGCGCTTGGGGTTGAAATTTTCAACATGATGACAGTTTATGTTCCCTTGCGCT	1440
QY	1441	GATGGCACAGATTCACAGGAAAAAGCCGGGCTTGCGCTATCAACAAGATCGCATGGCGCA	1500
Db	1441	GATGGCACAGATTCACAGGAAAAAGCCGGGCTTGCGCTATCAACAAGATCGCATGGCGCA	1500
QY	1501	CATTTACCAAGATCAACTCGAATTGATGCGGATGTCAGCTAATAACGTAGACCATTCGC	1560
Db	1501	CATTTACCAAGATCAACTCGAATTGATGCGGATGTCAGCTAATAACGTAGACCATTCGC	1560
QY	1561	AGCGGCTCAGTTCCATATCAATCCAAAGATTAAAAACCTGAGATTGGAACTTGGGT	1620
Db	1561	AGCGGCTCAGTTCCATATCAATCCAAAGATTAAAAACCTGAGATTGGAACTTGGGT	1620
QY	1621	CACTTTGGATGATTTTGAAGTCAACGGGATATCAGTTCCAGATCTTAATTCACATCCGCTT	1680
Db	1621	CACTTTGGATGATTTTGAAGTCAACGGGATATCAGTTCCAGATCTTAATTCACATCCGCTT	1680
QY	1681	TTCAAGTCTAATCCCGTAATTCAGGCGGATATGAGTTGAGTTTATATAAAAAAGCTCC	1740
Db	1681	TTCAAGTCTAATCCCGTAATTCAGGCGGATATGAGTTGAGTTTATATAAAAAAGCTCC	1740
QY	1741	CGAAGGTGCGGAGCTTTTATATCAAGATGATGCTTTAAGCTTTAAGCGGTAGGGCAG	1800
Db	1741	CGAAGGTGCGGAGCTTTTATATCAAGATGATGCTTTAAGCTTTAAGCGGTAGGGCAG	1800
QY	1801	AATGCTGCGGGGATGACCAACACACCCCAATAGATCTGACACCATTTTGCT	1860
Db	1801	AATGCTGCGGGGATGACCAACACACCCCAATAGATCTGACACCATTTTGCT	1860
QY	1861	CTTCAAGGCCCAATTGATGAGTCAAGCACTTTAATAGCAGTTGCGGTAAAGAAAG	1920
Db	1861	CTTCAAGGCCCAATTGATGAGTCAAGCACTTTAATAGCAGTTGCGGTAAAGAAAG	1920
QY	1921	AATAACCGTAATCAAGACCGGTACATCAAGCTTTAAGTGCACACGCGCAATTTG	1980
Db	1921	AATAACCGTAATCAAGACCGGTACATCAAGCTTTAAGTGCACACGCGCAATTTG	1980
QY	1981	CAGAGCAAAACCGGCAAACTCACAGAGACACCGGTTGCGCGAGCAGAGCAGTAATACA	2040
Db	1981	CAGAGCAAAACCGGCAAACTCACAGAGACACCGGTTGCGCGAGCAGAGCAGTAATACA	2040
QY	2041	AGTCCAAATGTTCCACACTAAGTAATAGGTATGATTTACGCACTTTACAGCAGCCAGA	2100
Db	2041	AGTCCAAATGTTCCACACTAAGTAATAGGTATGATTTACGCACTTTACAGCAGCCAGA	2100
QY	2101	GCCCCACGAGAAACCAATTAGCTGTTGTGTCGATGAGATTGAACCTAATCCGTAC	2160
Db	2101	GCCCCACGAGAAACCAATTAGCTGTTGTGTCGATGAGATTGAACCTAATCCGTAC	2160
QY	2161	CACGTATCTGAAACAATACCGGTATGGGCTTCGGGCAATGCCATTTTGTAAATTTTC	2220
Db	2161	CACGTATCTGAAACAATACCGGTATGGGCTTCGGGCAATGCCATTTTGTAAATTTTC	2220

Db 2161 CACTGTAAGCAATACCGTGTAGTGGGCTCGGCAATCGCATTTTGTAGAAATTC 2220

Qy 2221 ACCTGGCGGGCCCAACCATCAACTCTTCATCAAGTTTGGCCATGACCGTAAATGGCGAGAA 2280

Db 2221 ACCTGGCGGGCCCAACCATCAACTCTTCATCAAGTTTGGCCATGACCGTAAATGGCGAGAA 2280

Qy 2281 AATGTGCGCAATACCCCAATACGATTAAGTGGACACCCGAAAGTATTAATCCCAATACGA 2340

Db 2281 AATGTGCGCAATACCCCAATACGATTAAGTGGACACCCGAAAGTATTAATCCCAATACGA 2340

Qy 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATGTGTTTGAATAATCAAAACC 2400

Db 2341 AAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATGTGTTTGAATAATCAAAACC 2400

Qy 2401 TTTTCATGCTCATATTCGCGAGTCCGAGACAGGCGAAACGACGAGTGTGAGACTTTCTC 2460

Db 2401 TTTTCATGCTCATATTCGCGAGTCCGAGACAGGCGAAACGACGAGTGTGAGACTTTCTC 2460

Qy 2461 TAAATATGCCAATCATATTTCTAGAGGTAGAAAGATCAGACCGCGAAGATTTGAAAAA 2520

Db 2461 TAAATATGCCAATCATATTTCTAGAGGTAGAAAGATCAGACCGCGAAGATTTGAAAAA 2520

Qy 2521 ATGTGTGATGTGTGCTGCTGCGAAGACGCGGAAACTCTTTTACAGCGCATATGGCC 2580

Db 2521 ATGTGTGATGTGTGCTGCTGCGAAGACGCGGAAACTCTTTTACAGCGCATATGGCC 2580

Qy 2581 AAGGCTGACGAGATTTGCTGACAGTATACCAATATTTGGACCCCATCCATAGGAAT 2640

Db 2581 AAGGCTGACGAGATTTGCTGACAGTATACCAATATTTGGACCCCATCCATAGGAAT 2640

Qy 2641 CGCGGTTTCAACCGGTAACCCACGCGACGACGACCAATATAGAGTACCGTGTCT 2700

Db 2641 CGCGGTTTCAACCGGTAACCCACGCGACGACGACCAATATAGAGTACCGTGTCT 2700

Qy 2701 TGAAGATTGATATGATGCGCTTGCCTAAACCAATCATCATCTGCAATTTGGTGGGA 2760

Db 2701 TGAAGATTGATATGATGCGCTTGCCTAAACCAATCATCATCTGCAATTTGGTGGGA 2760

Qy 2761 AGCAATTCCTAAATAGAACTTTGGCTGATGCGCGGTTGCCATTTAAACCGCTGCGAC 2820

Db 2761 AGCAATTCCTAAATAGAACTTTGGCTGATGCGCGGTTGCCATTTAAACCGCTGCGAC 2820

Qy 2821 CATCGGCACTGCAAGAGTATGTAATATACAGATGAAAGCAATTTGCCCACTAGGCC 2880

Db 2821 CATCGGCACTGCAAGAGTATGTAATATACAGATGAAAGCAATTTGCCCACTAGGCC 2880

Qy 2881 TTTCTGTGTACGCGAAATCGGCGCTGCAG 2909

Db 2881 TTTCTGTGTACGCGAAATCGGCGCTGCAG 2909

RESULT 3
AE004153 9931 bp DNA linear BCT 10-OCT-2003
LOCUS
DEFINITION
Vibrio cholerae O1 biovar eltor str. N16961 chromosome I, section 61 of 251 of the complete chromosome.
ACCESSION
AE004153 AE003852
VERSION
AE004153.1 GI:9655103
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae O1 biovar eltor str. N16961
Vibrio cholerae O1 biovar eltor str. N16961
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 9931)
Heidelberg, J.F., Bisen, J.A., Nelson, W.C., Clayton, R.A., Gwim, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bae, S., Qin, H., Driscoll, I., Sellers, P., McDonald, L., Uitterlinden, T., Fleischmann, R.D., Nierman, W.C., and White, O.
DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae
JOURNAL
Nature 406 (6795), 477-483 (2000)

MEDLINE
20406833
PUBMED
10952301
REFERENCE
2 (bases 1 to 9931)
AUTHORS
Heidelberg, J.F., Bisen, J.A., Nelson, W.C., Clayton, R.A., Gwim, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bae, S., Qin, H., Driscoll, I., Sellers, P., McDonald, L., Uitterlinden, T., Fleischmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
TITLE
Direct Submission
JOURNAL
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
FEATURES
source
1..9931
/organism="Vibrio cholerae O1 biovar eltor str. N16961"
/mol_type="genomic DNA"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:243277"
/chromosome="I"
/note="biotype: El Tor"
288..602
/gene="VC0666"
288..602
/gene="VC0666"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF93831.1"
/db_xref="GI:9655104"
/translation="MSSDPSLSIVLPSKTTGENTMROSIVQLAVILVRLREER VKRKVRSSHYLPMNNPHLRLDGLGTFDGRFISLPAVVAEVRVRIIRVLAARI PT"
complement (640..1698)
/gene="VC0667"
complement (640..1698)
/gene="VC0667"
/note="similar to PID:882727 GB:U00096 PID:1789199 PID:2342573; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="oxido-reductase Tas, aldo/keto reductase family"
/protein_id="AAP93832.1"
/db_xref="GI:9655105"
/translation="MALRQGSAMQYTKLPHSSLEISIKI CLGTMTEGSONSQAFOQL DYALRGVNFIDTAEMVPPPTAQOQKTEBFGIMLAKSGRERI VLATKAGPRNV PYIRDMALDHEINHOAVDDSLRLQOTYIDLYOHMPOROTNGOLNVPYDQOE VTLITLRLANDLVLMGKVRVYGVNENPDGMSVYLRLAKREHLPRIVSIOMNYLIN RSFVGLAEISHLBEVYKILAYSPLAFGLSGKYLNGARPAARCTLHORFSRYPTFEG ILATEAYVALAQFGIDPQOMALAFVNDPFAVSNILGATVMEQLKSNLSDISLNA ELLQKIOBIGITVSNPCP"
complement (1891..2556)
/gene="VC0668"
complement (1891..2556)
/gene="VC0668"
/note="similar to GB:Y00113 SP:P06722 GB:U16361 PID:42065 PID:577425; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA mismatch repair protein Muth"
/protein_id="AAP93833.1"
/db_xref="GI:9655106"
/translation="MKPAPPTQOELLTRAQIAGUSFAELADEAGTVPPLDKDKGW VQLLEMLHAGASGRPOODFEHLIEKSIPISTYGRPLETTFCVAPLGVHGLTW EBSHVRNKLSTKYLMI PVQGEREIPLAECVSGPLMSPSPEBAOLKADWEELMGIW IGVKQIOTAKHSEVQLRPKAANGALTBAYGANORPIKALPRGYLATQTFQAQILOR YTA"
complement (2943..3035)
/gene="VC0669"
complement (2943..3035)
/gene="VC0669"
/gene="VC0669"

Db 7254 AACGATGTCCTTGGGCTTTTGTATTCCTTAATGCTGCGCCACTGCGCGGCATCTTC 7313
Qy 541 ACAGCTTATGAAATTCGCTTAGAAGCGTGTCTGTTCTTATTTCTTATTTAGTTTAT 600
Db 7314 ACAGCTTATGAAATTCGCTTAGAAGCGTGTCTTATTTCTTATTTCTTATTTAGTTTAT 7373
Qy 601 TGGTAAACCTGTCGCTAGGACGCGTATCCGACCTGTTTTAGCTGGATACGGTACAT 660
Db 7374 TGGTAAACCTGTCGCTAGGACGCGTATCCGACCTGTTTTAGCTGGATACGGTACAT 7433
Qy 661 CCGCTTCCTGTTGAAATACCTCGGTAGACGAGATGCTCACTTGGGTGTTGGTGGCTT 720
Db 7434 CCGCTTCCTGTTGAAATACCTCGGTAGACGAGATGCTCACTTGGGTGTTGGTGGCTT 7493
Qy 721 CATTTCAATGGGCAAAATCCTCTCTTAACTATGATGATCATCGGTATTTGATGATGAT 780
Db 7494 CATTTCAATGGGCAAAATCCTCTCTTAACTATGATGATCATCGGTATTTGATGATGAT 7553
Qy 781 TTGGTCTTACAGCGCGGTTGTATCAAGACGCTGTAGCAGCAAAATAGGGTATGATGAT 840
Db 7554 TTGGTCTTACAGCGCGGTTGTATCAAGACGCTGTAGCAGCAAAATAGGGTATGATGAT 7613
Qy 841 GAAACAGTATTTGATCTTTGTCAAGCGCATGCTGATCAAGGTGTTGGTGGTGAATGA 900
Db 7614 GAAACAGTATTTGATCTTTGTCAAGCGCATGCTGATCAAGGTGTTGGTGGTGAATGA 7673
Qy 901 AGCAACGGGCAAGCGTTGTTGACTGTGATTAATGCGATTTGACCTTACGATGAGTGGCA 960
Db 7674 AGCAACGGGCAAGCGTTGTTGACTGTGATTAATGCGATTTGACCTTACGATGAGTGGCA 7733
Qy 961 CAATCAGTTTCTCTAGTGACTACAGCAGAGTTTTTGGAAAGCTGCGCTAGCCGAGTT 1020
Db 7734 CAATCAGTTTCTCTAGTGACTACAGCAGAGTTTTTGGAAAGCTGCGCTAGCCGAGTT 7793
Qy 1021 GCTCGCTATATTCGTGTGTTGATTAATGCGCGGATTTTGGCAATTTAGTGAATCAAC 1080
Db 7794 GCTCGCTATATTCGTGTGTTGATTAATGCGCGGATTTTGGCAATTTAGTGAATCAAC 7853
Qy 1081 CTGGAGTGTATGCAATTTTAAACCAAGCATGAGCTCAACATCTTACCGTAAAGTGA 1140
Db 7854 CTGGAGTGTATGCAATTTTAAACCAAGCATGAGCTCAACATCTTACCGTAAAGTGA 7913
Qy 1141 GATGACATGAGGACGCGTGTATGTTTCAAGGTTAGAGCTTGGGCTAGCCGATGATG 1200
Db 7914 GATGACATGAGGACGCGTGTATGTTTCAAGGTTAGAGCTTGGGCTAGCCGATGATG 7973
Qy 1201 TCATATTTGACAGTGTAAAGATGTTGATGATTTGAGCGCTGCGTGTATGACCGAG 1260
Db 7974 TCATATTTGACAGTGTAAAGATGTTGATGATTTGAGCGCTGCGTGTATGACCGAG 8033
Qy 1261 TGAATTTCTTAACTTTCAACAATCCGGGTGAATTTCAATGAGGTTTGGCCCTTGCA 1320
Db 8034 TGAATTTCTTAACTTTCAACAATCCGGGTGAATTTCAATGAGGTTTGGCCCTTGCA 8093
Qy 1321 GTPACGCCATCATTTTTCATGCTGAGGAGTACCTTGTATCTCAACAGTATCAGGCTTC 1380
Db 8094 GTPACGCCATCATTTTTCATGCTGAGGAGTACCTTGTATCTCAACAGTATCAGGCTTC 8153
Qy 1381 ATGTGATGTGCTTGGGTTGAATTTCAACATGTGACAGTTTATGTGTTCTTGGCT 1440
Db 8154 ATGTGATGTGCTTGGGTTGAATTTCAACATGTGACAGTTTATGTGTTCTTGGCT 8213
Qy 1441 GATGACAGTATCAAGGGAAGGCGGGCTTGGGTTATCAACAAGTGTCAATGCGCA 1500
Db 8214 GATGACAGTATCAAGGGAAGGCGGGCTTGGGTTATCAACAAGTGTCAATGCGCA 8273
Qy 1501 CATTTACAGATCACTGAAATTTGAGCGGATGCTGAGCTTAAAGTGAAGCATTTCC 1560
Db 8274 CATTTACAGATCACTGAAATTTGAGCGGATGCTGAGCTTAAAGTGAAGCATTTCC 8333
Qy 1561 AGGCGCTCAGTTTCATATCAATCAAAAGATTAAACATGACAGATTTGGAACTTGGGT 1620
Db 8334 AGGCGCTCAGTTTCATATCAATCAAAAGATTAAACATGACAGATTTGGAACTTGGGT 8393

Qy 1621 CACTTTGATGATTTTGAAGTCAACCGGATATCAGTTTCAAGATCTTATTTCAATACCGGT 1680
Db 8394 CACTTTGATGATTTTGAAGTCAACCGGATATCAGTTTCAAGATCTTATTTCAATACCGGT 8453
Qy 1681 TTCAGTCTAATCCCGATTTACGCGGTTATGCTTGAATGGGTTTTATATTAAGGCTCC 1740
Db 8454 TTCAGTCTAATCCCGATTTACGCGGTTATGCTTGAATGGGTTTTATATTAAGGCTCC 8513
Qy 1741 CGAAGTCCGGAGCTTTTATTAAGATGATGCTTAAAGCTTAAAGCGTTAGCGGTTAG 1800
Db 8514 CGAAGTCCGGAGCTTTTATTAAGATGATGCTTAAAGCTTAAAGCGTTAGCGGTTAG 8573
Qy 1801 AATGCTCGCGGATGACGCAAAACACCCATATAGTAACTCAACCACTTTTGGCT 1860
Db 8574 AATGCTCGCGGATGACGCAAAACACCCATATAGTAACTCAACCACTTTTGGCT 8633
Qy 1861 CTTACAGCCCAAGTTGAGATGAGCTCAGCACTTTAATAGGAGTTTGGTAAAGAAAG 1920
Db 8634 CTTACAGCCCAAGTTGAGATGAGCTCAGCACTTTAATAGGAGTTTGGTAAAGAAAG 8693
Qy 1921 AATACCGTAAATCAAGACCGTATGACCTCAAGTTAAAGCTTAAAGTGAACCAAGCAATT 1980
Db 8694 AATACCGTAAATCAAGACCGTATGACCTCAAGTTAAAGCTTAAAGTGAACCAAGCAATT 8753
Qy 1981 CAGAGCAACACCGGCAACCTCACAGAGACGCGGTTGCGGCGAGAGAGCAAGTAAATAC 2040
Db 8754 CAGAGCAACACCGGCAACCTCACAGAGACGCGGTTGCGGCGAGAGAGCAAGTAAATAC 8813
Qy 2041 AGTGCCAAATGTTGCACTTAAGGTAAATGGTATGATTTCAAGCACTTTGACGCGCAG 2100
Db 8814 AGTGCCAAATGTTGCACTTAAGGTAAATGGTATGATTTCAAGCACTTTGACGCGCAG 8873
Qy 2101 GCCCAAGAGAGAACATTAAGTGTGTGTGTGATGAATGAATGAATCAATACCGTAA 2160
Db 8874 GCCCAAGAGAGAACATTAAGTGTGTGTGTGATGAATGAATGAATCAATACCGTAA 8933
Qy 2161 CACTGTACCTGAAGCAATACCGGTATGAGGCTCGGCGCAATCGCATTTTGTGATTTTC 2220
Db 8934 CACTGTACCTGAAGCAATACCGGTATGAGGCTCGGCGCAATCGCATTTTGTGATTTTC 8993
Qy 2221 AGGTGCGCGGCAACCATCAAACTTTCATCAAGTTTCCCATCAACCGTATGCGAGCA 2280
Db 8994 AGGTGCGCGGCAACCATCAAACTTTCATCAAGTTTCCCATCAACCGTATGCGAGCA 9053
Qy 2281 AATGTGCAATACCAATACGATTAAGTGCAGACCAACGAAAGTTTACCAATACCA 2340
Db 9054 AATGTGCGGATACCAATACGATTAAGTGCAGACCAACGAAAGTTTACCAATACCA 9113
Qy 2341 AAGCTGGGTTTCAAGCCCTGTATGACAGTTTGTGAATCGGTTGATTAATCAAAAC 2400
Db 9114 AAGCTGGGTTTCAAGCCCTGTATGACAGTTTGTGAATCGGTTGATTAATCAAAAC 9173
Qy 2401 TTTTCAATCATATCGCAGTGCAGAGCAGAGCGCAAGCAAGCTGTGAGACTTCTC 2460
Db 9174 TTTTCAATCATATCGCAGTGCAGAGCAGAGCGCAAGCAAGCTGTGAGACTTCTC 9233
Qy 2461 TAAATGCAAAATCATATTTTGAAGTGAAGATCAGCACCGCGAGAAATGAAAA 2520
Db 9234 TAAATGCAAAATCATATTTTGAAGTGAAGATCAGCACCGCGAGAAATGAAAA 9293
Qy 2521 AATGCGATGATGAGCACTGCGCAAGGCAAGGCAAACTTTCTTTAAGCGGATATGGCC 2580
Db 9294 AATGCGATGATGAGCACTGCGCAAGGCAAGGCAAACTTTCTTTAAGCGGATATGGCC 9353
Qy 2581 AAGGCTGAAGAGTATGTTGATCAAGTATGATCAATATTTGCAACCATAGAGAT 2640
Db 9354 AAGGCTGAAGAGTATGTTGATCAAGTATGATCAATATTTGCAACCATAGAGAT 9413
Qy 2641 CGCGGTTTCAACCGGTAACCAACCGGCAACGAGACCAACATTAATGAAGTCAACGCTGT 2700
Db 9414 CGCGGTTTCAACCGGTAACCAACCGGCAACGAGACCAACATTAATGAAGTCAACGCTGT 9473

Qy 2701 TGAGATTGAATCAGTCCGTTGGCATTAACCAATCATCAATCCTGCAATTGGGTGCGA 2760
Db 9474 TGAGATTGAATCAGTCCGTTGGCATTAACCAATCATCAATCCTGCAATTGGGTGCGA 9533
Qy 2761 AGCAATTCAATTAACCTTTGGCTTGATCGCGGTTGGCCCTTTAAACCGCTGCCAGC 2820
Db 9534 AGCAATTCAATTAACCTTTGGCTTGATCGCGGTTGGCCCTTTAAACCGCTGCCAGC 9593
Qy 2821 CATGCGCATGCGAAGATGTAATATACGATGAAGCAAGTTGGCCCAAGTAGGCC 2880
Db 9594 CATGCGCATGCGAAGATGTAATATACGATGAAGCAAGTTGGCCCAAGTAGGCC 9653
Qy 2881 TTTGCTGTGTCAGCGAATCGCGCTGCAG 2909
Db 9654 TTTGCTGTGTCAGCGAATCGCGCTGCAG 9682

RESULT 4
LOCUS VCTHYAG 1410 bp DNA linear BCT 01-NOV-2000
DEFINITION Vibrio cholerae thya gene.
ACCESSION Y17135.1 GI:3114980
VERSION Y17135.1 GI:3114980
KEYWORDS thya gene; thymidilate synthase.
SOURCE Vibrio cholerae
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.

REFERENCE 1 Valle, E., Ledon, T., Cedre, B., Campos, J., Valmaseda, T.,
Rodriguez, B., Garcia, L., Martero, K., Benitez, J., Rodriguez, S. and
Fando, R.
TITLE Construction and characterization of a nonproliferative El Tor
cholera vaccine candidate derived from strain 638
INFECTION Immun. 68 (11), 6411-6418 (2000)
JOURNAL 20490577
MEDLINE 11035753
PUBMED 2 (bases 1 to 1410)
AUTHORS Benitez, J. A.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1998) J. A. Benitez, Centro Nacional de
Investigaciones, Cientificas, PO Box 6990, La Habana, CUBA
FEATURES
source location/Qualifiers

1. 1410
/organism="Vibrio cholerae"
/mol_type="genomic DNA"
/strain="C7258"
/db_xref="taxon:666"
289. 1140
/gene="thya"
289. 1140
/gene="thya"
/codon_start=1
/transl_table=11
/product="thymidilate synthase"
/protein_id="CAA76645.1"
/db_xref="GI:3114981"
/db_xref="GDA:066108"
/db_xref="SWISS-PROT:066108"
translation="MRQYLDLCQRIVDGQWVENERTGRCLTVINADLTVDVGNNOF
PLVTRKSFVKAVALLEGLIRGYDNADPFLQRTKTDWADANLQAMLNRYGSGED
DMGRVYGVQGRAMAKPDGGHIDOLKIVDDLSRGVDGSEILNFPNPGSPHMGCLRPG
MYSHFSLIGDTLYLNTSTORSCDPLGINFMNVQYVPLALMAOITGKKPGIAVHKIV
NAHYDODELRMDVQLKREPPAPQFHINPKITLQDLFTWTTDDPDVTSYQYHPD
IQYPSV"

ORIGIN
Query Match 47.8%; Score 1390.8; DB 1; Length 1410;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1398; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 551 GAATTCGCTTGAAGAGCGGTGCTGTTCTTATATCTTAATTTGTTATTTGTTAACT 610
Db 1081 GATTTGACGTCACCGGATATCAGTTCCACGATCTTATTAATACCGTTTCAGTCTTA 1140

Db 1 GAATTCGCTTGAAGAGCGGTGCTTCTATTTCTTATTTGTTATTTGTTAACT 60
Qy 611 CGTCGCTAGGACAGCGATACCGACTGTTTATTTAGCTGATACCGTATTCGCTTCTT 670
Db 61 CGTCGCTAGGACAGCGATACCGACTGTTTATTTAGCTGATACCGTATTCGCTTCTT 120
Qy 671 GTGGAATACGTCGCTGAGCCAGATGCTCAGTTGGGCTGTTTGGTGGCTTATTTCAATG 730
Db 121 GTGGAATACGTCGCTGAGCCAGATGCTCAGTTGGGCTGTTTGGTGGCTTATTTCAATG 180
Qy 731 GGGCAATCCCTCCTTACCTATGATGATCATCGGTAATTTGATGATGTTGGCTTAC 790
Db 181 GGGCAATCCCTCCTTACCTATGATGATCATCGGTAATTTGATGATGTTGGCTTAC 240
Qy 791 AAGCGCGTTTGTATCAAGACCGTGTAGCAGCAAAATAGGTAAGTTAGTGAACAGTAT 850
Db 241 AAGCGCGTTTGTATCAAGACCGTGTAGCAGCAAAATAGGTAAGTTAGTGAACAGTAT 300
Qy 851 TTGATCTTTGTACAGCGATCGTGATCAAGGTGTTGGTTGAAATGAACGAGGC 910
Db 301 TTGATCTTTGTACAGCGATCGTGATCAAGGTGTTGGTTGAAATGAACGAGGC 360
Qy 911 AAGCGTTGTTGACTGTGATTAATGCCGATTTGACCTAAGATGTGGGCAACATCAGTTT 970
Db 361 AAGCGTTGTTGACTGTGATTAATGCCGATTTGACCTAAGATGTGGGCAACATCAGTTT 420
Qy 971 CCTAGTGACTACACGCAAGATTTTGAAGACCTGCGTAGCGAGTTCTCGGCTAT 1030
Db 421 CCTAGTGACTACACGCAAGATTTTGAAGACCTGCGTAGCGAGTTCTCGGCTAT 480
Qy 1031 ATTGCTGTATCAGATATGCGCGGATTTTGCATTAAGTATGATACCAAACTGGGATGCT 1090
Db 481 ATTGCTGTATCAGATATGCGCGGATTTTGCATTAAGTATGATACCAAACTGGGATGCT 540
Qy 1091 AATGCCATTTTAACCAAGCATGCTCAACAACTCTTAACGTTAAAGATGATGACATG 1150
Db 541 AATGCCATTTTAACCAAGCATGCTCAACAACTCTTAACGTTAAAGATGATGACATG 600
Qy 1151 GAGCGCGTATGAGTGTTCAGGGTAGAGCTTGGGCTAAGCCTGATGTGTATATTTGAC 1210
Db 601 GAGCGCGTATGAGTGTTCAGGGTAGAGCTTGGGCTAAGCCTGATGTGTATATTTGAC 660
Qy 1211 CAGTTGAAAAGATTTGATGATTTGAGCCGTGCGTTGATGACCGAGTGAATTTCTT 1270
Db 661 CAGTTGAAAAGATTTGATGATTTGAGCCGTGCGTTGATGACCGAGTGAATTTCTT 720
Qy 1271 AACTTCTACATCCGGGTGAATTTTCACATGGGGTGTGGCCCTTGTGATGATGACCAT 1330
Db 721 AACTTCTACATCCGGGTGAATTTTCACATGGGGTGTGGCCCTTGTGATGATGACCAT 780
Qy 1331 CATTTTCATGCTGGGGGATACCTTGTATCTCAACAGTACTAGCGTTTCATGATGTG 1390
Db 781 CATTTTCATGCTGGGGGATACCTTGTATCTCAACAGTACTAGCGTTTCATGATGTG 840
Qy 1391 CCTTGGGGTGAATTTTCAACATGCTGAGGTTTATGTGTTCTTTGCGCTGATGCGACAG 1450
Db 841 CCTTGGGGTGAATTTTCAACATGCTGAGGTTTATGTGTTCTTTGCGCTGATGCGACAG 900
Qy 1451 ATCACAGGAAAAGCCGGGCTTGGCGTATCAAGATGCTGATGCGGCAATTTACCA 1510
Db 901 ATCACAGGAAAAGCCGGGCTTGGCGTATCAAGATGCTGATGCGGCAATTTACCA 960
Qy 1511 GATCAACTGGAATGATCGGATGCTGAGGCTTAAACGTAAGCCATTTCCAGCGGCTGAG 1570
Db 961 GATCAACTGGAATGATCGGATGCTGAGGCTTAAACGTAAGCCATTTCCAGCGGCTGAG 1020
Qy 1571 TTCCATATCAATCCAAAGATTTAAACACTGACGATTTGGAACCTTGGGTCACTTTGAT 1630
Db 1021 TTCCATATCAATCCAAAGATTTAAACACTGACGATTTGGAACCTTGGGTCACTTTGAT 1080
Qy 1631 GATTTGACGTCACCGGATATCAGTTCCACGATCTTATTAATACCGTTTCAGTCTTA 1690
Db 1081 GATTTGACGTCACCGGATATCAGTTCCACGATCTTATTAATACCGTTTCAGTCTTA 1140

QY 1691 TCCCGTATTGAGCGGTATGCTTGATGCTTTTATATATAAAAAAGCTCCGGAAGTCCG 1750
 Db 1141 TCCCGTATTGAGCGGTATGCTTGATGCTTTTATATATAAAAAAGCTCCGGAAGTCCG 1200
 QY 1751 GAGCTTTTATATACAGATGATGCTTTAAAGCTTAAGCGGTTAGGGCAAGATGCTGCCG 1810
 Db 1201 GAGCTTTTATATACAGATGATGCTTTAAAGCTTAAGCGGTTAGGGCAAGATGCTGCCG 1260
 QY 1811 GGGATAGCACAACACACCCCAATAAGTAATCTACACCAACATTTTGCTCTTACAGAGC 1870
 Db 1261 GGGATAGCACAACACACCCCAATAAGTAATCTACACCAACATTTTGCTCTTACAGAGC 1320
 QY 1871 CAAGTTAGATGAGCTCAGACACCTTTAATAGGCAAGTTCGGTAAGAAAGAAATACCGTAA 1930
 Db 1321 CAAGTTAGATGAGCTCAGACACCTTTAATAGGCAAGTTCGGTAAGAAAGAAATACCGTAA 1380
 QY 1931 ATCAAGACCGGTAGCCATCAAGTTAAAGCTT 1960
 Db 1381 ATCAAGACCGGTAGCCATCAAGTTAAAGCTT 1410

RESULT 5
 AX009483 1222 bp DNA linear PAT 06-SEP-2000
 LOCUS AX009483
 DEFINITION Sequence 3 from Patent WO961634.
 ACCESSION AX009483
 VERSION AX009483.1 GI:9996768
 KEYWORDS
 SOURCE Vibrio cholerae
 ORGANISM Vibrio cholerae
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.

REFERENCE
 1 Carlin, N. and Lebens, M.R.
 Method of producing thy a<->strains of vibrio cholerae, such strains and their use
 Patent: WO 961634-A 3 02-DEC-1999;
 JOURNAL CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
 FEATURES
 source 1..1222
 Location/Qualifiers
 1..1222
 /organism="Vibrio cholerae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:666"

ORIGIN
 Query Match 42.0%; Score 1222; DB 6; Length 1222;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1688 TAATCCCGTATCAGCGGTATGCTTGATGCTTTTATATATAAAAAAGCTCCGGAAGT 1747
 Db 1 TAATCCCGTATCAGCGGTATGCTTGATGCTTTTATATATAAAAAAGCTCCGGAAGT 60
 QY 1748 CGGAGCTTTTATACAGATGATGCTTTAAAGCTTAAGCGGTTAGGGCAAGATGCTG 1807
 Db 61 CGGAGCTTTTATACAGATGATGCTTTAAAGCTTAAGCGGTTAGGGCAAGATGCTG 120
 QY 1808 CCGGGAGTAGCACAACACACCCCAATAAGTAATCTACACCAACATTTTGCTTTACAA 1867
 Db 121 CCGGGAGTAGCACAACACACCCCAATAAGTAATCTACACCAACATTTTGCTTTACAA 180
 QY 1868 GCCCAAGTTGAGATAGCTCAGACCTTTAATAGGCAAGTTCGGTAAGAAAGAAATACCG 1927
 Db 181 GCCCAAGTTGAGATAGCTCAGACCTTTAATAGGCAAGTTCGGTAAGAAAGAAATACCG 240
 QY 1928 TAAATCAAGACCGGTAGCCATCAAGTTAAAGCTTAAGTGCAGCAGCGCAATTTGCAAGCA 1987
 Db 241 TAAATCAAGACCGGTAGCCATCAAGTTAAAGCTTAAGTGCAGCAGCGCAATTTGCAAGCA 300
 QY 1988 AACACGGCAAACTACACAGAGA CAGCGGTTCGGCGAGCAGACAGTAATACAGTGCCA 2047
 Db 301 AACACGGCAAACTACACAGAGA CAGCGGTTCGGCGAGCAGACAGTAATACAGTGCCA 360

QY 2048 ATGTCGACCTTAAGTAATGAGTAGATTTTCACGCACTTTGACGACCGCAGAGCCACG 2107
 Db 361 ATGTCGACCTTAAGTAATGAGTAGATTTTCACGCACTTTGACGACCGCAGAGCCACG 420
 QY 2108 AAGAGAACCTTAAGCTGCTGTTGTTGCTGATGTAAGATTAACCTAATACCTGTA 2167
 Db 421 AAGAGAACCTTAAGCTGCTGTTGTTGCTGATGTAAGATTAACCTAATACCTGTA 480
 QY 2168 CCGTAAGCAATACCGGTATGAGGAGCTGCGCAATGCAATTTTGTAATTTCAAGTGG 2227
 Db 481 CCGTAAGCAATACCGGTATGAGGAGCTGCGCAATGCAATTTTGTAATTTCAAGTGG 540
 QY 2228 CCGCAACCATCAAACTCTTCATCAAGTTGCCATCAACGTAATGCGACGAAATGCTC 2287
 Db 541 CCGCAACCATCAAACTCTTCATCAAGTTGCCATCAACGTAATGCGACGAAATGCTC 600
 QY 2288 GCAATACCAATACGATTAAGTGCACACCAACGAAAGTAATTAACCAATACGAAAGCTG 2347
 Db 601 GCAATACCAATACGATTAAGTGCACACCAACGAAAGTAATTAACCAATACGAAAGCTG 660
 QY 2348 GTTTCAGCGCTGTGATGACAGGTTGTAATCGGTTGTAATTAATCAAAACCTTGATG 2407
 Db 661 GTTTCAGCGCTGTGATGACAGGTTGTAATCGGTTGTAATTAATCAAAACCTTGATG 720
 QY 2408 CTCATATCGCAGTCGCAAGCAGAGGCGCAACGACGAGTGAGACTTTCTTAAATG 2467
 Db 721 CTCATATCGCAGTCGCAAGCAGAGGCGCAACGACGAGTGAGACTTTCTTAAATG 780
 QY 2468 CCAAAATCATTTCTAGAGTAGAAGATACGACCGCGAAGATTTGAAAAATCGTGG 2527
 Db 781 CCAAAATCATTTCTAGAGTAGAAGATACGACCGCGAAGATTTGAAAAATCGTGG 840
 QY 2528 ATGCTGCGACTGCGCAAGCAACGCGGAATCTTTCTTACAGCGCATATGCGCAAGCTG 2587
 Db 841 ATGCTGCGACTGCGCAAGCAACGCGGAATCTTTCTTACAGCGCATATGCGCAAGCTG 900
 QY 2588 ACGAGATTTGCTGACAGTAGTACCAATATTGGCAACCATACCATGAGTAAGATTCGGTT 2647
 Db 901 ACGAGATTTGCTGACAGTAGTACCAATATTGGCAACCATACCATGAGTAAGATTCGGTT 960
 QY 2648 TCAACCGGTAAACCCACCGGCAACGAGACCAACATTAATAGAGTACCGTCTTGAGAT 2707
 Db 961 TCAACCGGTAAACCCACCGGCAACGAGACCAACATTAATAGAGTACCGTCTTGAGAT 1020
 QY 2708 TCAATCAGTCCGCTTCCCATTAACCATCATCAATCTGCAATGGTGGGAAACCAAT 2767
 Db 1021 TCAATCAGTCCGCTTCCCATTAACCATCATCAATCTGCAATGGTGGGAAACCAAT 1080
 QY 2768 TCAATCAGTCCGCTTCCCATTAACCATCATCAATCTGCAATGGTGGGAAACCAAT 2827
 Db 1081 TCAATCAGTCCGCTTCCCATTAACCATCATCAATCTGCAATGGTGGGAAACCAAT 1140
 QY 2828 ACTGCAAGAGTAGTAATTAACGATGAAGCAAGTTCGCCAAGTATGAGCTTTGCTG 2887
 Db 1141 ACTGCAAGAGTAGTAATTAACGATGAAGCAAGTTCGCCAAGTATGAGCTTTGCTG 1200
 QY 2888 GTCAGCGAAATCGCGCTGCGAG 2909
 Db 1201 GTCAGCGAAATCGCGCTGCGAG 1222

RESULT 6
 AY143429 1236 bp DNA linear BCT 28-SEP-2002
 LOCUS AY143429
 DEFINITION Vibrio cholerae mutant chymidilase synthetase (thyA) gene, complete cds.
 ACCESSION AY143429
 VERSION AY143429.1 GI:23343945
 KEYWORDS
 ORGANISM Vibrio cholerae
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

REFERENCE
1 (bases 1 to 1236)
Xia,X.
JOURNAL
Submitted (20-AUG-2002) Microbiology, Institute for Infectious
Disease Control and Prevention, Chinese Center for Disease Control
and Prevention, P.O. Box 5, Changping, Beijing 102206, P.R. China
FEATURES
Location/Qualifiers

source
1. 1236
/organism="Vibrio cholerae"
/mol_type="genomic DNA"
/db_xref="taxon:666"
248..1099
/gene="chxA"
/note="chxA"
/note="does not catalyze the conversion of dUMP into dTMP"
/codon_start=1
/transl_table=11
/product="mutant thymidilate synthetase"
/protein_id="AA016904.1"
/db_xref="GI:23343946"
/translation="MRQYLDLCQRIVDGQWVENERTGRCLTYINADUTYDVGNNQF
PLVTRKSFVKAABVLELGLIRGYDNADFRQLGRTKTDANANLQAMLPNRYRGE
DMGRVYGVQGRAMAPDGGHIDOLKRIVDLSRGVDGEILNPNRPFHMGCLRP
MYSHFSLIGDTLYNSTORSQDPLMLNFMWYVYFLAAMAOTGKPKGLAYKIV
NAHIYODLEIMRDVQLKREPAPQPHINPKIKTLQDLFTVTLDDVDVGYQPHDP
I0YPSV"

ORIGIN

Query Match 41.8%; Score 1216.8; DB 1; Length 1236;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1224; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

592 TTGTTTATTTGTTAACTGTCCTGCTAGGACGCGTATCCGAGCTGTTTAACTGATGA 651
1 TTGTTTATTTGTTAACTGTCCTGCTAGGACGCGTATCCGAGCTGTTTAACTGATGA 60
652 CGGTACATCCGCTCTCTGTTGAAATAGCTCCGTAGACAGATGCTCAGCTGGTCTGT 711
61 CGGTACATCCGCTCTCTGTTGAAATAGCTCCGTAGACAGATGCTCAGCTGGTCTGT 120
712 TGTGGCTTCATTTCAATGAGGCGAAATCCTCTCCTTACCTATGATGATCGATTTT 771
121 TGTGGCTTCATTTCAATGAGGCGAAATCCTCTCCTTACCTATGATGATCGATTTT 180
772 GATGATGTTTGTGCTTTACAGCGCGGTTTGTATCAAGCGGTGATGACGCAAAATAGG 831
181 GATGATGTTTGTGCTTTACAGCGCGGTTTGTATCAAGCGGTGATGACGCAAAATAGG 240
832 TAGTTAGTGAACAGTATTTAGATTTTGTCAAGCGCATGTCATCAAGTGTGGT 891
241 TAGTTAGTGAACAGTATTTAGATTTTGTCAAGCGCATGTCATCAAGTGTGGT 300
892 TGAATAATGAACAGCGGCAAGCGGTTGTTGACTGTGATTAATGCCGATTTGACCTACA 951
301 TGAATAATGAACAGCGGCAAGCGGTTGTTGACTGTGATTAATGCCGATTTGACCTACA 360
952 TGTGGCAACAATCAGTTTCTCTAGTGACTACAGCAAGAGTTTGGAAAGTCCGCT 1011
361 TGTGGCAACAATCAGTTTCTCTAGTGACTACAGCAAGAGTTTGGAAAGTCCGCT 420
1012 AGCCGAGTTGCTGGCTATATTCGTGTTTACGATTAATGCGGCGGATTTTCCGCAATTA 1071
421 GGCCTGAGTTGCTGGCTATATTCGTGTTTACGATTAATGCGGCGGATTTTCCGCAATTA 480
1072 TACCAAAACCTGGAGTGAATGCAATTAACCAAGATGAGCTCAACATCTTAACG 1131
481 TACCAAAACCTGGAGTGAATGCAATTAACCAAGATGAGCTCAACATCTTAACG 540
1132 TAAAGGTGAGATGATGAGGACGCGTGTATGCTGTTGAGGATGAGCTTGGGCTTAAGC 1191

541 TAAAGGTGAGATGATGAGGACGCGTGTATGAGCTTCAAGGTTAGAGCTTGGGCTTAAGC 600
1192 TGAATGTTGATATTTAGCAAGTGAAGAAAGATTTGTTGATTTGAGCCGTGGCGTTGA 1251
601 TGAATGTTGATATTTAGCAAGTGAAGAAAGATTTGTTGATTTGAGCCGTGGCGTTGA 660
1252 TGACCGAGGTGAATTTCTTAATCTTACATATCCGGGTGAATTTTCAATGGGGTGTTCG 1311
661 TGACCGAGGTGAATTTCTTAATCTTACATATCCGGGTGAATTTTCAATGGGGTGTTCG 720
1312 CCTTGCATGTTACAGGCAATCTTTTTCATGCTGGGGGATACCTTGTATCTCAACAGTAC 1371
721 CCTTGCATGTTACAGGCAATCTTTTTCATGCTGGGGGATACCTTGTATCTCAACAGTAC 780
1372 TCAGGCTTCATGATATGCTTGGGCTGAATTTCAACATGATGACAGTTTATGTT 1431
781 TCAGGCTTCATGATATGCTTGGGCTGAATTTCAACATGATGACAGTTTATGTT 840
1432 CTTGCGCTGATGGACAGATCAAGGAAAGCGGCGCTTGGCGTATCAACAGATGT 1491
841 CTTGCGCTGATGGACAGATCAAGGAAAGCGGCGCTTGGCGTATCAACAGATGT 900
1492 CAATGCGGACATTTTACCAAGATCACTGCAATGATGAGGAGTGGACAGCTTAAACGTA 1551
901 CAATGCGGACATTTTACCAAGATCACTGCAATGATGAGGAGTGGACAGCTTAAACGTA 960
1552 GCCATTTCCACGCGCTCAGTTTCAATATCAATCAAAAGATTAAACACTGACAGATTGGA 1611
961 GCCATTTCCACGCGCTCAGTTTCAATATCAATCAAAAGATTAAACACTGACAGATTGGA 1020
1612 AACTTGGGTCATCTTGTGATGATTTTGAATGATGATGATGATGATGATGATGATGAT 1671
1021 AACTTGGGTCATCTTGTGATGATTTTGAATGATGATGATGATGATGATGATGATGAT 1080
1672 ATACCCGTTTCAAGTCTATACCCGATTCAGCGGATGATGATGATGATGATGATGATGAT 1731
1081 ATACCCGTTTCAAGTCTATACCCGATTCAGCGGATGATGATGATGATGATGATGATGAT 1140
1732 AAAAGCTCCGCAAGGTGCGGAGCTTTTATACAGATGATGATGATGATGATGATGATGAT 1791
1141 AAAAGCTCCGCAAGGTGCGGAGCTTTTATACAGATGATGATGATGATGATGATGATGAT 1200
1792 TAGGGCAAGATGCTGCGGCGGATGACGCAAAAC 1827
1201 TAGGGCAAGATGCTGCGGCGGATGACGCAAAAC 1236

RESULT 7
VCAJ10968/c 1591 bp DNA linear BCT 21-SEP-1998
LOCUS VCAJ10968
DEFINITION Vibrio cholerae nphA gene.
ACCESSION AJ010968
VERSION AJ010968.1 GI:3646476
KEYWORDS nphA gene; phosphate pump.
SOURCE
ORGANISM
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1
Lebens,M., Soderlund,L.O., Lundquist,P. and Carlén,N.I.A.
A putative sodium dependent phosphate pump in Vibrio cholerae
JOURNAL
Unpublished
2 (bases 1 to 1591)
AUTHORS
Carlén,N.I.A.
TITLE
Direct Submission
COMMENT
Submitted (17-SEP-1998) Carlén,N.I.A., SBL Vaccin AB, Department of
Molecular biology, FMB, Jv 30, Stockholm, S-105 21, SWEDEN
nphA gene is located inbetween the nhar gene (AJ002395) and the
chxA gene
(AJ006514) both running in opposite directions.
FEATURES
source
1. 1591
/organism="Vibrio cholerae"

```

/mol_type="genomic DNA"
/strain="JS1569"
/db_xref="taxon:666"
245..1588
/gene="nptCA"
-35_signal
245..250
/gene="nptCA"
-10_signal
267..272
/gene="nptCA"
373..376
/gene="nptCA"
385..1532
/gene="nptCA"
CDS
/codon_start=1
/cranal_table=11
/product="sodium dependent phosphate pump"
/protein_id="CAA09443.1"
/db_xref="GI:3646477"
/db_xref="GOA:O87918"
/db_xref="SPTREMBL:O87918"
/translation="MINOATSPAPISLTTLGLRMNLAFMLYLLLAAMVSGGPKM
ATGDOKVLEEFASHPILAGIMIGLVATLILIOSSYTSIIYGLVAGGLPVEPIPMV
GANITGVNTLVLSGMRCKEFPRAFSATIHDPNLAVALIPIEMREILKRV
SHLVSPILATGDMSKGFDPIKPIKPVITGLEQLSVLNGFVALIVIGIATIF
VALTWGKLMKSLMGRAREILQNALGRPLNGIAGVTVVLVIOSSSTTSLMVDLV
GSGVLKREIYPTLGNIGTCTTALLAATAVSGEPAVFAVLAQIALVHLSFNMAATVLI
YGIPLREIPIKAEILISTWACKSKMVVSYLLGVFVVIPIGSLALTA"
1550..1588
stem_loop
/gene="nptCA"
/note="shared with thyA gene running in opposite
direction"
/function="transcription termination"

```

ORIGIN

```

Query Match      40.8%; Score 1187; DB 1; Length 1591;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1723 TTATATAAAAAAGCTCCGAGAGTGGGAGCTTTTATATACAGATGATGCTTTAAAGC 1782
DB 1591 TTATATAAAAAAGCTCCGAGAGTGGGAGCTTTTATATACAGATGATGCTTTAAAGC 1532
QY 1783 TTAAGCGTTAGGGCAAGATGCTGCCGGGATGACGACAAACACCCATTAAGTAAGT 1842
DB 1531 TTAAGCGTTAGGGCAAGATGCTGCCGGGATGACGACAAACACCCATTAAGTAAGT 1472
QY 1843 CACCAACCACTTTTCTCTTCAAGACCCCAAGTTGAGATGAGTCAAGACCTTTAATAG 1902
DB 1471 CACCAACCACTTTTCTCTTCAAGACCCCAAGTTGAGATGAGTCAAGACCTTTAATAG 1412
QY 1903 CAGTTGGCTTAAGAAAGAAATACCGTAATCAAGACCGTAATCAAGTTAAAGCTTAA 1962
DB 1411 CAGTTGGCTTAAGAAAGAAATACCGTAATCAAGACCGTAATCAAGTTAAAGCTTAA 1352
QY 1963 GTGCACCAAGGCAATTTGACAGCAAAACGCAAACTCACGAGACAGCGGTTGCGG 2022
DB 1351 GTGCACCAAGGCAATTTGACAGCAAAACGCAAACTCACGAGACAGCGGTTGCGG 1292
QY 2023 GAGCAGAGCAGTAATACAAAGTGCATGTTCCGACCTTAAGTAATGGGTAGATTTCAG 2082
DB 1291 GAGCAGAGCAGTAATACAAAGTGCATGTTCCGACCTTAAGTAATGGGTAGATTTCAG 1232
QY 2083 CACTTTCAGCAGCCGACGAGCCACGAGAGAAACATTAAGGCTGTGTGTGCATGAAGA 2142
DB 1231 CACTTTCAGCAGCCGACGAGCCACGAGAGAAACATTAAGGCTGTGTGTGCATGAAGA 1172
QY 2143 TTGAACTAATACCGTAACCACTGTACTGTGAAGCAATACGTTAGTGGGCTGCGGCAAT 2202
DB 1171 TTGAACTAATACCGTAACCACTGTACTGTGAAGCAATACGTTAGTGGGCTGCGGCAAT 1112
QY 2203 CGCATTTTGTAGATTTCAGTGTGGGCGGCAACATCAACTCTTCATCAGTTTGGCCAT 2262
DB 1111 CGCATTTTGTAGATTTCAGTGTGGGCGGCAACATCAACTCTTCATCAGTTTGGCCAT 1052

```

```

QY 2263 CACCGTAATGCGCAGCAAAAATGCTCGCAATACCCAAATACGATTAAGTGCACACCAAGAA 2322
DB 1051 CACCGTAATGCGCAGCAAAAATGCTCGCAATACCCAAATACGATTAAGTGCACACCAAGAA 992
QY 2323 AGTATTACCAATATACGAAAGCTGGGTTTCAGACCCCTGTGATGACAGGTTTGGTAATCG 2382
DB 991 AGTATTACCAATATACGAAAGCTGGGTTTCAGACCCCTGTGATGACAGGTTTGGTAATCG 932
QY 2383 TTTGATTAATAATCAAAACCTTTCATGCTCATATGCGCAGTCCGAGAGAGCGGCAACGAG 2442
DB 931 TTTGATTAATAATCAAAACCTTTCATGCTCATATGCGCAGTCCGAGAGAGCGGCAACGAG 872
QY 2443 CCAGTGTAGACTTCTCTTAAATATGCCAAACATCATTTGTAGAGTAGAAGATACAGAC 2502
DB 871 CCAGTGTAGACTTCTCTTAAATATGCCAAACATCATTTGTAGAGTAGAAGATACAGAC 812
QY 2503 CCGGAGAGATTTGAAAAAATCGTGATGTCGTCGCAACCTGCGGAAAGCAACGCGAAATCTTC 2562
DB 811 CCGGAGAGATTTGAAAAAATCGTGATGTCGTCGCAACCTGCGGAAAGCAACGCGCGAAATCTTC 752
QY 2563 TTTACAGCCCATATGCGCAGAGCTGACGAGATTTGTCACAGTAGTACCAATATTGCG 2622
DB 751 TTTACAGCCCATATGCGCAGAGCTGACGAGATTTGTCACAGTAGTACCAATATTGCG 692
QY 2623 ACCCATCAACATATGAAATGCGGTTTCAACCGGTAAACCAACGCGCAACGACCAACAT 2682
DB 691 ACCCATCAACATATGAAATGCGGTTTCAACCGGTAAACCAACGCGCAACGACCAACAT 632
QY 2683 AATGAAAGTCAACCGTCTTGAAGATTTGAATCAAGTGCCTGTCACATCAATATCA 2742
DB 631 AATGAAAGTCAACCGTCTTGAAGATTTGAATCAAGTGCCTGTCACATCAATATCA 572
QY 2743 TCTGCAATTTGGTGGGAGCAATTTCAATTAAGATTTGGCTTGAATGCCCGTTGCCA 2802
DB 571 TCTGCAATTTGGTGGGAGCAATTTCAATTAAGATTTGGCTTGAATGCCCGTTGCCA 512
QY 2803 TTTAAACCGCTGCCGACCATGCGGCTGCAAGAAAGTAGTAATACAGATGAAGACCA 2862
DB 511 TTTAAACCGCTGCCGACCATGCGGCTGCAAGAAAGTAGTAATACAGATGAAGACCA 452
QY 2863 GTTTCGCAACGTAGGCTTTCGTGTGTCAGCAAAATCGCGCTGCAG 2909
DB 451 GTTTCGCAACGTAGGCTTTCGTGTGTCAGCAAAATCGCGCTGCAG 405

RESULT 8
AP005074 298900 bp DNA linear BCT 05-MAR-2003
LOCUS Vibrio parahaemolyticus DNA, chromosome 1, complete sequence, 2/11.
DEFINITION AP005074.BA000031
ACCESSION AP005074.1 GI:28805287
VERSION
KEYWORDS
SOURCE
ORGANISM
Vibrio parahaemolyticus
Vibrio parahaemolyticus
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1
AUTHORS Nasu, H., Iida, T., Sugahara, T., Yamachi, Y., Park, K. S., Yokoyama, K.,
Makino, K., Shingawa, H. and Honda, T.
TITLE A filamentous phage associated with recent pandemic Vibrio
paraahaemolyticus O3:k6 strains
JOURNAL J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
MEDLINE 20295086
PUBMED 10834969
REFERENCE
2
AUTHORS Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T.,
Tagomori, K., Iijima, Y., Najiima, M., Nakano M., Yamashita, A.,
Kubota, Y., Kimura, S., Yasunaga, T., Honda, T., Shingawa, H.,
Hattori, M. and Iida, T.
TITLE Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
distinct from that of V cholerae
JOURNAL Lancet 361 (9359), 743-749 (2003)

```

MEDLINE 22508454
PUBMED 12620739
REFERENCE 3 (bases 1 to 298900)
AUTHORS Oshima,K., Kurokawa,K., Makino,K., Yokoyama,K., Yasunaga,T.,
Honda,T., Shinagawa,H., Hattori,M. and Iida,T.
TITLE Direct Submission
Submitted (09-Apr-2002) Ken Kurokawa, Osaka University, Genome
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:ken@gen-info.osaka-u.ac.jp,
URL: http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
Fax:81-6-6879-2047)
COMMENT
This clone was isolated from a patient presenting with acute
gastroenteritis.
FEATURES
source
1..298900
/organism="Hibrio parahaemolyticus"
/mol_type="genomic DNA"
/strain="O3:K6"
/sub_strain="RIMD 2210633"
/db_xref="taxon:670"
/chromosome="1"
115..315
/gene="VP0305"
115..315
/gene="VP0305"
/note="similar to PIR:S56442 percent identity 49 in 63 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC58568.1"
/db_xref="GI:28805288"
/translation="MLRPAVYRPLQVAFVKTLLKGFIVGVSGFCFQDNGKVLPLD
VKDQRLAVFKVKNQAISSLSA"
complement(399..1037)
/gene="VP0306"
complement(399..1037)
/gene="VP0306"
/note="similar to GB:AAF95689.1 (AE004324) percent
identity 74 in 208 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="peptide methionine sulfoxide reductase"
/protein_id="BAC58569.1"
/db_xref="GI:28805289"
/translation="MLNKQTLISIDALPGRQPMQIDCHFWNQSLTAPLAHQQQ
ILGWCQFAGARLFWQLDGVSTSVYAGGFTPNPYEAVCTGKGTGVVAVFDE
RVLISLAQLAVFWEKHDPQGRQGNLGTYSALITYSDQCEIADRSKIYQOAL
BAELRSTITTEIVPAGPYFAETVHQYLAKNPDGVCIGTGVCPPSLQ9"
1205..2917
/gene="VP0307"
1205..2917
/gene="VP0307"
/note="similar to GB:AAF95689.1 (AE004323) percent
identity 66 in 571 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC58570.1"
/db_xref="GI:28805290"
/translation="MIKKALPTIIGLVAVSOSAVADVSLKLGIDGALDENVKAYLSS
IPEKYSTSLRQARLDOSTEALNALGYFAKITYSISEGNDLIVIHKLVPYKIK
VMDVVISGEAKDEEPTNLIAKSPKVRILNQGYSILKSGIRNALAQGVFNDPK
LAKLEIVIELNEANVRLHYDSGIRYHFGVEELPTSGIWMENRVSMRPEISGPIYSD
VGEYQNLSTNDPSSVYEPDLSKLEGRLEPLIYSLAPAKNOIFETGIGYSTDGV
RGTLLKMKPWSVARGHSFNTALSLSPGQITTAGYKPILDVLRVYOLQRLKHLND
RDTSELSNLAVERHWLTDGWHKTVYRHLIYENFSQGLQDDGQPVLPQATFESTRY
RGGSPFMWGDQSVTVEYGDPLVLETRVLRILGKSSWIRIGENHRGLFRLEGANI
TEEFKLSPSLAFAGDNNIRGYEESISPVDEGALTGAYTILISTLEYOYRYGN
WMAATFYDIDAFNDTPWKSGAGVIRWASPVGVSFDFAWGLDEKNNRPIHFSL

gene
CDS
GPEL"
2914..6681
/gene="VP0308"
2914..6681
/gene="VP0308"
/note="similar to GB:AAF95688.1 (AE004323) percent
identity 54 in 1258 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC58571.1"
/db_xref="GI:28805291"
/translation="MIKKRMYKMWLSLSLGLLLITFYATVLTFTHPGLKALWGA
EKALPOLKVENIQSLPFRPALQVSVESLSIADVERIALAINRCFDPKVCVD
ELAQGVNFMPELPPASEEPEETPLRSISTVPIIFVNVSFNDINVNLGNQIDW
OTFSTALSMOGRVIAPTALKDINVLAPAPAEBAEAEPEKTDATPODIVPEW
IPLVEARLDIHNFKLGERPVIVNHLGIARAGRVDTKLELMEPEKGLSTQ
VTLSDYPIKQADALYKQADAKQKSLASGSGVCDLSNATISLVQNAIKGDIQD
LKTQLPPVSLIKVQAQWPLFGQSDYGVSVPSLAKQSLDGEVALFTKSGKDIYAV
DVALNGKLTLEQIDLESIVVKTLEGBLSGKVMAMAPINWQADLNKNIQPLQWKE
AEGDISGSLSTSGSLTEQGWVSLPKLIDIGIRGPLVNEGQLEASDKKGEIDIC
LITQGLALSHGNOLSAKXIDKQIINDVENPDPFAKSVYDLGAKQKQALRGSAL
EPDINIDLAINQIMWQQAIVETTLTGQVTPLPAPRANSLIANNITVDIKYDSAD
LEVSGDEKTLQIIDVSVLSLEIBGTFKQKPEMTWDGALRRLTSSQGPWSIQ
KSTAVKVNIDKQIANVOAHCWLQAKSSVCILTEDISVKTQAKLAINDFNDQKQPL
PPEKTLQSVVAAQFAKMAPEKKEVTLSTVMPGQVEALQEPISKVWMSFSPKAL
AKDLKLEWLPDVNDGDLGKVSILNVSSEKPTIDKVALSTPLDLPLIGESYL
FKAMNTDIALSGVDMHPKVGQPLIDOMKLGQVTPIDINSQGVNLFKQHOADLA
GITTPDKLEITGPADMRDLODMRTTRARVAKELKUMPMWVKIKVSPNDITVTPOL
AKVGNINLPEWRIVIELEPSSAVGSDVITLAKDLOPVDVAMAPPNVETDINIK
GDDPOLAFLGKGLKGLKSLANTYQDKPEFIVGEVNIYDGSYRSFGQDIVLEGKILMN
GADQDPYVSIKAIINPNTODDVIAGRVGEPADPSVLEIFSEPMQANALSYLRG
ODIDGESGNAMTTLLIGLSLAKSGRVGGEIGEAFGVODIQLDPAGSGDSQVYSGV
ILPGLQYXGVGINSIGSEFTVRYRIMQDLYLBAVSGVDSAVDLLYQPEPD"
6726..7073
/gene="VP0309"
6726..7073
/gene="VP0309"
/note="similar to GB:AAF95687.1 (AE004323) percent
identity 64 in 115 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC58572.1"
/db_xref="GI:28805292"
/translation="MQHLIFPYGTLRGEYVNHHTLSSAQPLGHETDAQSFYNVGSY
PALSQGSISIGEVYLLDDDTLVALDKEDVPVYRBSIATPFGQAMITINYQTEOL
TEETASGWCQRV"
7170..7532
/gene="VP0310"
7170..7532
/gene="VP0310"
/note="unknown"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC58573.1"
/db_xref="GI:28805293"
/translation="MRVLIIGAMMMLISACAQTLTPASTVYTDWQAFGQALDLARE
LSORIAKLDGTNATPELIMAYONGYQEGKQFECQSAVMLGVVGRPVFGICDDVDP
FROHDYAGRSSAGAP1"
complement(7607..8137)
/gene="VP0311"
complement(7607..8137)
/gene="VP0311"
/note="similar to GB:AAF95686.1 (AE004323) percent
identity 90 in 176 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11

Query Match	Best Local Similarity	Matches 1278;	Conservative	Score 1035;	DB 1;	Length 301442;
				Mismatches 405;	Indels 0;	Gaps 0;
Qy	8	TTTGTATGCTCAGGGTATCTGCAGTTTCCCAATATTTGACCCGCTATGTTTGCATC	67			
Db	215617	TTCAATTATGCTCAGGGTATCTGCAGTTTCCCAATATTTGACCCGCTATGTTTGCATC	215558			
Qy	68	GGCCCTCAGCGGTCGCTGCTGATAGCTGATATTTGGTGGGTTTCCCTTTTGGCTATG	127			
Db	215557	GGACCCATCTCAGTTTCGTTGGTGAACGGTTGATATCTATCTATTTGGTGGTGGATG	215498			

QY	128	TGGTGGCAATCGCCGAGCGGATCGCGGGCAGAGGTGTGACGCGGTAGCAAGCTCT	187
Db	215457	TGGCTCGCCAATCGCCGAGCAGACAACCGGGCAGTGGCTGACGGAGACAAAGTTTCG	215438
QY	188	GACTTGTTATTCGCGGCTTTTAAAGTGTAGTATCGGTGGCCGAGTGGTATGTGATC	247
Db	215437	GACTTGCTGTTGCGTGGATTCCTTAAGCGTGTGTGGGTGTGTCGCAATTGGGGTACGTGCTG	215378
QY	248	TTCTCAATTTTGAATCTGTTCCCTTGACCCCTCTTATTTATTCAAAGTGTGACCTGGC	307
Db	215377	TTCTATTAACCTTCGATCTCTCTCTGCTGATCCGATTTTACCTCTTTAAAGCTGTGACAGGT	215318
QY	308	GGCATGTCCTTCCACAGCGCGCTTATTTGAGTGTATCACCGCATGTTCTGTATGTGCGCT	367
Db	215317	GGCATGTCCTTCCACAGCGCGCTTGTGCTGGGTGTCTTACGGCTATAGCTCTGTATCGGGAAG	215258
QY	368	AAAAACAACGCACTTCTTTGTGTGTGCGCATTTTGTGTCCCTTTAAGTGCATTCGCT	427
Db	215257	AAAAATGGCCGAACCTTTTGTGGCGGTGGGAGCACTTGTCCTCGTTAAGCAATTTGGC	215198
QY	428	TTGGGAGATGGGACGATCGGTAACTTTATGAATATGAACCTTTGGGACGAGTAACGAT	487
Db	215137	TTAGGGGATAGGGCGCTTAGGCAACTTTATGAACGGAGCTTTGGGGCAGATGACCGAT	215133
QY	488	GTGCTTGGGCTTTTGTATTCCTTAATGTGTGCCACATGCGCGCCATCTTTCAGACTT	547
Db	215137	GTGCTTGGGCGATGTGTGTCCCAACGGCGGACCACTACAGCCACCTTGCAATTG	215078
QY	548	TATGAATTCGCTTAGAAGAGCGTGTCTGTCTTTATCTTAATGTGTATATGTATTA	607
Db	215077	TATGAATGTGCATTAAGAAAGGGTGTGTCTGTCTTTATCTCAACGTGTTTATCCGCAAA	215018
QY	608	CCTCGTCGCTAGCAGCGGATCCCGACCTGTTTATAGCTGATACCGTACATTCGCTTC	667
Db	215017	CCAACCTTTATAGATCCGATCTGTGTTATTCCTTGCTGTGTATGTATATTCGCTTTC	214958
QY	668	CTTGTGAATACGTCCGTGAGCCAGATGCTCATGTTGGTCTGTGTGGCTTCATTTCA	727
Db	214957	CTGTGTGAATACGTCCGTGAGCCAGATGCGCACTGGGCTGTGTGGTGTTCATTTCA	214898
QY	728	ATGGGGCAAAATCCTGCTTACCTTAACCTTAAGGATCATCGATTTTGAATGATGTGGTCT	787
Db	214897	ATGGGACAGATCTTGTGCTCACCGATGATTAATGTGTGATTTGGCCTGATGGCTTGGCT	214833
QY	788	TACAAGCGCGTTTGTATCAAGACCGGTGAGCAGAAAATAGGGTATTAAGTGAACAG	847
Db	214837	TATAAACGCGCACTATCATGAACAAGTAACGTGAATGAAGTAATGAAGTAACAA	214778
QY	848	TATTTTAGATCTTTGTACAGCGCATCTGTCATCAAGGTGTTTGGGTGAAAATGAACGACG	907
Db	214777	TATTTTAGAATGTGCGGTGCGATCTGTATGAAGAGGCACTGGGTGTAAAAATGAACGTACC	214718
QY	908	GCGAAGCGTGTGTACCTGTGATTAATGCGCAATTTGACCTAGATGTGGCAACAATCAG	967
Db	214717	GCGAAGCGCTGCTTACTGTATGATCAACGGGATTTTAACCTAGATGTTGCCAACAATCAG	214658
QY	968	TTTCTCTAGTGACTACAGCAGAAAGTTTGGAAAAGCTGCCGTAGCCGAGTTGCTCGGC	1027
Db	214657	TTTCTCTTGTGTTACCAACCGTAAAGTGTTCGTGAAGGCGCGCGTGTGCAAGTTCCTTGGC	214598
QY	1028	TATATTCGTTGATGATTAATGCGCGGATTTTGGCAATTAAGTATCCAAAACCTTGGGAT	1087
Db	214597	TACATTCGATGATTAATGATTAATGCGGAAGTTTTGGCAAGTATGATTCGAAAACCTTGGGAT	214538
QY	1088	GCTAATGCAATTTAAACCAAGCATGCTCAACAAATCCTTAACGTTAAAGGTGAGATGAC	1147
Db	214537	GCGAAGCGCAATCTCAATGATGCTGTTGAATTAACCAATATGCAAAAGCGCAAGATGAC	214478
QY	1148	ATGGGACGCGTATGATGATGATTCAGGGTATGAGTGTGGCTAAACCTGATGTGTCTATTT	1207
Db	214477	ATGGGCGGTGTATGCGGTATTAAGGTGAGTGTGGCGCAACCTGATGTGTGTCTATATTC	214418

QY 1208 GACCAATTGATTGATGATTTGACCGCGGTGATGACCGAGGTGAAATT 1267
 |||||
 Db 214417 GACCAATTGCGTAAATTGTCATGACCTGACCGCGGAGTGAAGATCGCGGAGATC 214358
 |||||
 QY 1268 CTTACTTCTCAAAATCCGGGTAAATTTCAATGGGGTGTTCGGCCCTTGATGATGAC 1327
 |||||
 Db 214357 CTCAATTTCTCAACCCAGGTAAATTTCAATAGGGTGTTCGGCCCTTGATGATGAC 214298
 |||||
 QY 1328 CATCATTTTCTGCTGGGGGATACCTGTATCTCAACAGTACTCAGCGTTCATGTAT 1387
 |||||
 Db 214297 CATCATTTCTGCTGGGGGATACCTGTATCTCAACAGTACTCAGCGTTCATGTAT 214238
 |||||
 QY 1388 GTGCCCTTGGGGTGAATTTCAACATGTGACGGTTTATGTTCCTTGGCGTGA 1447
 |||||
 Db 214237 GTGCCCTTGGGGTGAATTTCAACATGTGACGGTTTATGTTCCTTGGCGTGA 214178
 |||||
 QY 1448 CAGATCAAGGAAAAAGCGGGCTTGGGTATCACAAGATGTCAATGCCACATTTAC 1507
 |||||
 Db 214177 CAGATTCAGGGAAGAAACCGGGCCAGCATTCATAGATTTGAATGACATATCTAT 214118
 |||||
 QY 1508 CAAGATCACTGAAATGATGGCGATGTCAGCTAAACGTCAGCATTCACGCGCT 1567
 |||||
 Db 214117 CAAGATCACTGATGATGTCAGTATGTGCACTCAACCGAAGACCTTACAGCGCA 214058
 |||||
 QY 1568 CAGTTCATATCATTCCAAAAGATTAAACACTGACAGATTGGAACCTTGGGTCACTTG 1627
 |||||
 Db 214057 AGTTCCACATCAACCTGAGATCAATCTTTGGAAGATCTGGAACCTTGGGTCACTTG 213998
 |||||
 QY 1628 GATGATTTGACGTCAACCGATATCATGTTCCACAGATCTTATTCATACCGGTTTCAAGTC 1687
 |||||
 Db 213997 GATGATTTGACGTCAACCGATATCATGTTCCACAGATCTTATTCATACCGGTTTCAAGTC 213938
 |||||
 QY 1688 TAA 1690
 |||||
 Db 213937 TAA 213935

RESULT 10
 LOCUS AP005332 249150 bp DNA linear BCT 04-DEC-2003
 DEFINITION Vibrio vulnificus Y016 DNA, chromosome I, complete genome, section 3/14.
 ACCESSION AP005332 BA000037
 VERSION AP005332.1 GI:37197435
 KEYWORDS
 SOURCE
 ORGANISM
 . Vibrio vulnificus Y016
 Vibrio vulnificus Y016
 Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
 REFERENCE
 AUTHORS 1
 Chen, C.Y., Wu, K.M., Chang, Y.C., Chang, C.H., Tsai, H.C., Liao, T.L., Liu, Y.M., Chen, H.J., Shen, A.B., Li, J.C., Su, T.L., Shao, C.P., Lee, C.T., Hor, L.I., and Tsai, S.F.
 TITLE Comparative Genome Analysis of *Vibrio vulnificus*, a Marine Pathogen
 JOURNAL Genome Res. 13, 2577-2587 (2003)
 REFERENCE 2 (bases 1 to 249150)
 AUTHORS Chen, C.Y., Wu, K.M., and Tsai, S.F.
 TITLE Direct Submission
 JOURNAL Submitted (29-May-2002) Shih Peng Tsai, National Health Research Institutes, Division of Molecular and Genomic Medicine; 128, Yen-Chu-Yuan Road, Sec 2, Taipei, Taiwan 115, Republic of China (E-mail: pepsai@nhri.org.tw, Tel: 886-2-8146-1041, Fax: 886-2-2789-0484)
 COMMENT This sequence was determined by the Sequencing Core of the National Yang-Ming University Genome Research Center (YMGRC; http://genome.ym.edu.tw).
 FEATURES
 source 1. 249150
 location/Qualifiers
 /organism="Vibrio vulnificus Y016"
 /mol_type="genomic DNA"
 /strain="Y016"
 /db_xref="taxon:196600"
 /chromosome="I"

gene 299..979
 /gene="VV0511"
 CDS 299..979
 /gene="VV0511"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93275.1"
 /db_xref="GI:37197436"
 /transl_table=11
 /translation="MNVRSRKNNLRLTTKTPFGRLPIQTEHGPLVLDYLEKMENTID IALDEYRLMAIRVDLARPRLRKUENSGNWTPTLRLOQIDHSGRKRGSVRVH PCKVRIAWREBRSINDHHLVLPNKDRPNWLGKTQIQENIGSTVEKAVAGV DYEANGIVHPSKRNDSVVIHRLNCSDAFDADFEPKHISYLAKENKHPGSGR RNFSHSHK"
 gene 865..888
 /gene="rpt112"
 repeat_region 865..888
 /gene="rpt112"
 /note="2 units of 12 bp in-tandem repeat"
 /rpt_type=tandem
 /rpt_unit="TGATGCGTTGA"
 gene 1040..1708
 /gene="VV0512"
 CDS 1040..1708
 /gene="VV0512"
 /note="identified by GeneMark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93276.1"
 /db_xref="GI:37197437"
 /translation="MTQNLNVAHQRAHTTDRTPFNGKPLYPFDGLVLEYLEIDRV LTELANKPRIVAYAHNINLPSDFGVDLVFAHFPRLBETNELCKDYDTRVY OQVIVRYMTKASRSTSDHHLILLPDRLFQNGTREGGRFLYKVKVRAKMSVSE AYRQPCSGLGISTNBSFELSDAEVFPLOLNFPRISRLARPTLNDRKRSFG CSHD"
 gene complement(1286..1441)
 /gene="VV0513"
 CDS complement(1286..1441)
 /gene="VV0513"
 /note="identified by Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAC93277.1"
 /db_xref="GI:37197438"
 /translation="MKELPIKQDKMNVIAAGLTRLRDPIANRLVSIISGVVLT ELISLRP"
 gene 1733..2368
 /gene="VV0514"
 CDS 1733..2368
 /gene="VV0514"
 /note="identified by GeneMark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC93278.1"
 /db_xref="GI:37197439"
 /translation="MPSFTHKALLMANRTYLPSTIDRTFEDYPLYSKGLYESA LTVADVBEQALSQFDIALARLDLPEDHQDILINDYFNILKEKLDLSLFV WRKEDKVPNSHNYVLMFSDYQHGPAICWDKXNELVGHKTMQEAIEKRYSGEAR SLVFLDRNGYGVGTRCSKQVNIKNVTFHMSLSLAEMEKRYCFGSLE"
 gene 2436..2642
 /gene="VV0515"
 CDS 2436..2642
 /gene="VV0515"
 /note="identified by GeneMark and Glimmer2"
 /codon_start=1
 /transl_table=11
 /product="predicted transcriptional regulator"
 /protein_id="BAC93279.1"
 /db_xref="GI:37197440"

```

/translacion="MKMKRLKDKVMSLGLGSGTLYKFMADETDPKPSVPLGAAVA
WSESEIEWMEQDLARNPESQ"
/gene="VV0516"
/2725..3162
/gene="VV0516"
/2725..3162
/gene="VV0516"
/codon_start=1
/transl_table=11
/product="ribonuclease HI"
/protein_id="BAC93280.1"
/db_xref="GI:37197441"
/translacion="MSYSITYDGAAPRNQHGCTGGIGLVMEDENEYHESFTIR
KDNALDELALILEALEVADGVITYSDSYCVKGFNIWDDKRCMRADKKPVKN
RHLQOVDELSSKRYVEKKAHSGVEGNERADLLAVEAAL"
/3182..4423
/gene="VV0517"
/3182..4423
/gene="VV0517"
/3182..4423
/gene="VV0517"
/note="identified by Genemark and Glimmer2"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAC93281.1"
/db_xref="GI:37197442"
/translacion="MGCTPQAVILIHYYCWLREPLITPOYCPRNEISLMEKFRMP
VILGNPPLSPDEVSAGFLGYPSPREKIKMLASVKSLLPTDLINISKEGIRNK
STRQFAMOLARVPRKRGMSLNSLPGGIHKSLSFMSHITLKGVOGGSSTGEINL
GLVSTYIDRCVAYGQQLAFIQDAPINENQHELISGYLLPIPLDTLLSBOGKALYK
NMLKSPSELASKETBOAMCFCYDFWLSMSVIDVFLMEWDHVAEYTPSVAQOYG
LRQVNPBBERETLGRFGLIKQINQYLAJSEYVLLPVSEYERNGLLKRDVQAQF
KSWRSCKSLSPVSLSTFPANMDAGRGADLLIYALFMRALDKVGSRLPIDYTKVGS
RLSPDIYTSRSRYFOHYAP"
/4310..4381
/gene="rpt1173"
/4310..4381
/gene="rpt1173"
/note="2 units of 36 bp in-tandem repeat"
/rpt_type=tandem
/rpt_unit="AAGGTGGTACGCTCTCTACCTGATATCTACAG"
/complement(4874..5101)
/gene="VV0518"
/complement(4874..5101)
/gene="VV0518"
/note="identified by Genemark and Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC93282.1"
/db_xref="GI:37197443"
/translacion="MNKSTLLITATLVSTSSAFAGGTILGNSADSGAVNSVGCT
QYKMDYKADQAKENSTITVPKCEAIKEX"
/complement(5149..5508)
/gene="VV0519"
/complement(5149..5508)
/gene="VV0519"
/note="identified by Genemark and Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC93283.1"
/db_xref="GI:37197444"
/translacion="MNTLIRITMALLISITACSDQDAQVYTADEVISINPFWG
TILVFDIQTIDSTVITGVINRGNCLPAGTISLSRNVLAKFGETTYGYSNKCTVS
VREIVTSSVGVFYTF"
/complement(5532..5861)
/gene="VV0520"
/complement(5532..5861)
/gene="VV0520"
/note="identified by Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC93284.1"

```

```

/db_xref="GI:37197445"
/translacion="MALPTIREMSDEQRQVINAHTALLANMLKISSQVLOQEKALI
EPLGSLVQCDALLFLRLSGQTHIRTLALPLIDEASVSFTPTDFNVERFMEHA
BIRASOI"
/complement(5965..6273)
/gene="VV0521"
/complement(5965..6273)
/gene="VV0521"
/note="identified by Genemark and Glimmer2"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC93285.1"
/db_xref="GI:37197446"
/translacion="MSNKTSEIGLSALLHTLLMLIPQHWITVRELOQDALLIHRTT
RSIKRYLDIIVDFNVNECDSSMSPHVRYKTSBQLLKINKQEMALLATQVYLVTAGS"
/complement(6412..6588)
/gene="VV0522"
/complement(6412..6588)

```

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	Length	DB 3	Length
8	TTTGTATGCTTCAGGAGTATTCGACGATTTCCCAATATTCACCCGATATTTGATC	67	175147	TTCTATATGTCACAGGTTACCTGCTTTCTTCTAATATGACCCGCTTTTCTCGATA	175206	68	GGCCCTCTAGCGGCTGCTGATATGCTGATATTTGGTGGCTTCTTTTGTATG	127
175147	TTCTATATGTCACAGGTTACCTGCTTTCTTCTAATATGACCCGCTTTTCTCGATA	175206	175207	GGACCATCTCAGTTGTTGTTGATACGGGTTGATGATATTTGTTTGTGGATG	175266	128	TGCTTGGCAATTCGCGACCGATGCGGCGGAGTGTGATGACGGTGAAGCTCT	187
175267	TGCTTGGCAATTCGCGACCGATGCGGCGGAGTGTGATGACGGTGAAGCTCT	175326	175327	GATTTGCTTGTGCTGATTCCTTAGCGCTTGCTGAGTGTGCTGAGTGTGCTG	175386	188	GATTTGCTTGTGCTGATTCCTTAGCGCTTGCTGAGTGTGCTGAGTGTGCTG	247
175327	GATTTGCTTGTGCTGATTCCTTAGCGCTTGCTGAGTGTGCTGAGTGTGCTG	175386	248	TTCTAATATTTGATTCCTTCTTCTGACCCCTTTATTTATTCAGATGACCTGAC	307	175387	TTCTAATATTTGATTCCTTCTTCTGACCCCTTTATTTATTCAGATGACCTGAC	175446
175387	TTCTAATATTTGATTCCTTCTTCTGACCCCTTTATTTATTCAGATGACCTGAC	175446	308	GGCATCTCTTCCACAGCGGCTTATTTGAGTGTGATGACCGCATGTTCTGTATGCGCT	367	175447	GGCATCTCTTCCACAGCGGCTTATTTGAGTGTGATGACCGCATGTTCTGTATGCGCT	175506
175447	GGCATCTCTTCCACAGCGGCTTATTTGAGTGTGATGACCGCATGTTCTGTATGCGCT	175506	368	AAAAACCAACGACCTCTTGTGATGAGCGGCAATTTGTCCTTTAGTGCATTCGCT	427	175507	AAAAACCAACGACCTCTTGTGATGAGCGGCAATTTGTCCTTTAGTGCATTCGCT	175566
175507	AAAAACCAACGACCTCTTGTGATGAGCGGCAATTTGTCCTTTAGTGCATTCGCT	175566	428	TTGGGATGAGGACGATCGGTATCTTATTAAGTGAATTTGGGAGAGATGACGAT	487	175567	TTGGGATGAGGACGATCGGTATCTTATTAAGTGAATTTGGGAGAGATGACGAT	175626
175567	TTGGGATGAGGACGATCGGTATCTTATTAAGTGAATTTGGGAGAGATGACGAT	175626	488	GTGCTTGGGCTTTTGTATTCCTTAATGATGACCACTGCGCGCATCTTCAAGCTT	547	175627	GTGCTTGGGCTTTTGTATTCCTTAATGATGACCACTGCGCGCATCTTCAAGCTT	175686
175627	GTGCTTGGGCTTTTGTATTCCTTAATGATGACCACTGCGCGCATCTTCAAGCTT	175686	548	TATGATTCGCTTATGAGGAGCGTGTCTGTTCTTATTTCTTAATTTGTTATGTTAA	607	175687	TATGATTCGCTTATGAGGAGCGTGTCTGTTCTTATTTCTTAATTTGTTATGTTAA	175746
175687	TATGATTCGCTTATGAGGAGCGTGTCTGTTCTTATTTCTTAATTTGTTATGTTAA	175746	608	CCTGCTCGCTAGGAGCGATTCGGAATGTTTATGATGATGATGATGATGATGATGAT	667	175747	CCTGCTCGCTAGGAGCGATTCGGAATGTTTATGATGATGATGATGATGATGATGAT	175806
175747	CCTGCTCGCTAGGAGCGATTCGGAATGTTTATGATGATGATGATGATGATGATGAT	175806	668	CTTGGAATTAAGTCCGTGAGGCAAGTCTAGTGTGCTGTTGTTGTTGTTGTTGTTG	727	175807	CTTGGAATTAAGTCCGTGAGGCAAGTCTAGTGTGCTGTTGTTGTTGTTGTTGTTG	175866
175807	CTTGGAATTAAGTCCGTGAGGCAAGTCTAGTGTGCTGTTGTTGTTGTTGTTGTTG	175866						

```

QY 788 ATGGGCAATCTCTCTTACCTATGCTGATCATCGGATTTTGTATGATGTTGGTCT 787
Db 175867 ATGGGACGATCTTGTGTCACCGATGATTAATGGTGATTCGCCCTGATGCTGGGCT 175926
QY 788 TACAGCGCGGTTGATCAAGACCGGTGAGCAGCAAAATAGGATGTTAGGTGAACAG 847
Db 175927 TATTAACGTGGCCATATCAAGCAAAAGTACCGTGAAGTAAAGAAATGAAATGAAACAA 175986
QY 848 TATTTAGATCTTTGTACAGCGCATCGTCATCAAGGTGTTGGGTGAAATGAACGAG 907
Db 175987 TATTTAGATGTTGCGCGTCGATCGTGAATGAAGGCACTGGGTGAAATGAACGTACC 176046
QY 908 GCGAAGCGGTTGTTGATCTGTGATTAATGCCGATTTGACCTTACGATGTGGGCAACATCAG 967
Db 176047 GGTAAACCGTGCCTTACTGTGTGATCAACGCGGATTTAACTTAAAGATGTTGCCAACATCAG 176106
QY 968 TTTCCCTAGTACTACACGCAAGAGTTTGTGAAAGCTGCGGTGACGAGTTGCTCGGC 1027
Db 176107 TTTCCCTTGGTTTACACGCGTAAAGTTTCTGAAAGCGCGGTTGCAAGTTGCTTGGC 176166
QY 1028 TATATTCGTGTTAGCATATGCGCGGATTTTTCGCAATTTAGTACCAAACTCGGAT 1087
Db 176167 TACATTCGAGGCTATGACATCGGGAAGTTTTCGAAAGTTAGGTACGAAACCTGGAT 176226
QY 1088 GCTAATGCCAATTAAACCAAGCATGCTCAACATCTTACCGTAAAGGTAGCATGAC 1147
Db 176227 GCGAAGCGCAATCTCAATGATGCTTGTGTAATACCACTATTCGAAAGCGCATGATGAC 176286
QY 1148 AAGGACGCGGTATATGTTGATGAGGTGATGAGGTTGAGGCTGATGAGTGTGATAT 1207
Db 176287 AAGGCGCGGTCTACGCGCTACAGGTCGAGTCTGGCGCAAGCTGATGAGTGTGATATC 176346
QY 1208 GACGAGTTGAAAGATGTTGATGATTTGACCGCGGTGATGATGACCGAGTGAAT 1267
Db 176347 GACCAATTCGCAAAATGTCATGATGACGACGCGGATGATGATGCGGGGAGATC 176406
QY 1268 CTTAATCTTCAATCCGGGTGAATTTTCAATGAGGTTGTTGCGCTTGACATGACAC 1327
Db 176407 CTTAATCTTCAATCCAGGTGAATTTTCAATGAGGTTGTTGCGCTTGATGATGACAT 176466
QY 1328 CATCATTTTTCATGCTGGGGATACCTGTATCTCAACAGTACCAAGTTCATGAT 1387
Db 176467 CATCATTTTTCATGCTGGGGATACCTGTATCTCAACAGTACCAAGTTCATGAT 176526
QY 1388 GTGCCCTGGGTTGAATTTCAACATGTCAGGTTTATGTTCTTGGCTGATGGA 1447
Db 176527 GTACCGCTGGGCTGAATTTCAACATGTCAGGTTTATGTTCTTGGCTGATGGA 176586
QY 1448 CAGATCAAGGAAAAAGCCGGGCTTGGCGTATCAACAGATCTCATATGCGCATTTAC 1507
Db 176587 CAGATCAAGGAAAAAGCCGGGCTTGGCGTATCAACAGATCTCATATGCGCATTTAC 176646
QY 1508 CAAATCACTGATGATGTCGCGATGTCAGCTTAAACGTGACCATTTCCACGCGCT 1567
Db 176647 GAAAGTCACTTATGCACTGATGTCGCGATGTCAGCTTAAACGTGACCATTTCCACGCGCA 176706
QY 1568 CAGTTCATATCAATCCAAAGATTAACACTGACGATTTGGAACCTTGGCTCATTTG 1627
Db 176707 AGCTTCATATCAACCTGATGATCAATTTTGAAGATCTGGAACCTTGGCTCATTTG 176766
QY 1628 GATGATTTGACGTCACCGGATATCATGTTCCAGATCTTATTTCAATCCCGTTTCAATC 1687
Db 176767 GATGATTTGACGTCAGGATATGATGATGATGATGATGATGATGATGATGATGATGAT 176826
QY 1688 TAA 1690
Db 176827 TAA 176829

```

RESULT 11
 AX020429 852 bp DNA linear PAT 07-SEP-2000
 LOCUS AX020429
 DEFINITION Sequence 1 from Patent WO935271.

```

ACCESSION AX020429
VERSION AX020429.1 GI:10044145
KEYWORDS
SOURCE
ORGANISM Vibrio cholerae
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
REFERENCE
1 Campos,G.J., Ledon,P.T., Rodriguez,G.B., Benitez,R.J., Pando,C.R.,
Silva,C.A. and Valle,D.E.
Vibrio cholerae vaccine candidates and method of their constructing
Patent: WO 93/5271-A 1 15-Jul-1999;
JOURNAL
CAMPOS GOMEZ JAVIER (CU); LEDON PEREZ TALENA YAMILIE (CU); RODRIGUEZ
GONZALEZ BORIS LOTIS (CU); BENITEZ ROBLES JORGE ANTONIO (CU); CENT
NAC INVESTIG SCIENT (CU); PANDO CALADA RAFAEL ALFREDO (CU); SILVA
CABRERA ANISIA JUANA (CU); VALLE DIAZ EDGAR (CU)
location/Qualifiers
FEATURES
source
1..852
/organism="Vibrio cholerae"
/mol_type="unassigned DNA"
/db_xref="taxon:666"
1..852
/codon_start=1
/transl_table=1
/protein_id="CAC07770.1"
/db_xref="GI:10044146"
/db_xref="EMBL:CAC07770"
translation="MRQYLDCRIVQGVWVENERTGKRLATVINDLTVDGNNOF
PLATRKSPMKRAVAAYELIGYRGVDNAADPROGKTMDNANLQAMLNPKRGD
DMGRVYVQGRAMAKPDGSHIDQJLKIYVDLSRVDRLGELINFGEPFMGCLRPC
MYSHFSLIDPLTINSTQKSDVPLQKDFNFMVVYPLAMQITGKRPGLAYHKIV
NAHYVODLEIMRVDVQKRPFPAPQPHINKITLDLETWTLDDFDVDTGYQFHPD
IQYPSV"
misc_difference 1..3
/note="Xaa= Methionine"
ORIGIN
Query Match 28.9%; Score 841.4; DB 6; Length 852;
Best Local Similarity 99.3%; Pred. No. 1.7e-249;
Matches 845; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 840 TGAACAGTATTGATTTTGTGACGGCATCGTCAAGGTGTTGGTGAATAAG 899
Db 2 TGAGCAGTATTGATCTTGTGACGGCATCGTCAAGGTGTTGGTGAATAAG 61
QY 900 AACGAAGGGGAAGGTTGTACTGATTAAGCCGATTGACCTAGATGAGGCA 959
Db 62 AACGAAGGGGAAGGTTGTACTGATTAAGCCGATTGACCTAGATGAGGCA 121
QY 960 ACAATCACTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTGCCGTAGCCGAGT 1019
Db 960 ACAATCACTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTGCCGTAGCCGAGT 1019
QY 122 ACAATCACTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTGCCGTAGCCGAGT 181
Db 122 ACAATCACTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTGCCGTAGCCGAGT 181
QY 1020 TGCTCGGCTATATTTGTGTTACGATTAATGCGCGGATTTTGGCAATTAGGTACAAA 1079
Db 182 TGCTCGGCTATATTTGTGTTACGATTAATGCGCGGATTTTGGCAATTAGGTACAAA 241
QY 1080 CCGGAGTGTAAAGCCAAATTTAAACCAAGATGAGCTCAACATCTTACCGTAAAGGTG 1139
Db 242 CCGGAGTGTAAAGCCAAATTTAAACCAAGATGAGCTCAACATCTTACCGTAAAGGTG 301
QY 1140 AGGATGACATGGGAGCGCTGATGAGTTCAGGGTAGAGTTGGGCTAAGCTGATGGTG 1199
Db 302 AGGATGACATGGGAGCGCTGATGAGTTCAGGGTAGAGTTGGGCTAAGCTGATGGTG 361
QY 1200 GTCATATTTGACCGTTGAAAAAGATTTGTGATGATTTGAGCCGTGAGTATGACGAG 1259
Db 362 GTCATATTTGACCGTTGAAAAAGATTTGTGATGATTTGAGCCGTGAGTATGACGAG 421
QY 1260 GTGAAATCTTAACTTTCACAAATCCGGGTGAATTTTCAACATGGGCTGTTGGCCCTTGA 1319
Db 422 GTGAAATCTTAACTTTCACAAATCCGGGTGAATTTTCAACATGGGCTGTTGGCCCTTGA 481

```

QY 1320 TGTACAGCCATCTTTTTCATGCTGGGGGATACCTTGATCTCAACAGACTCAGGGT 1379
 DB 482 TGTACAGCCATCTTTTTCATGCTGGGGGATACCTTGATCTCAACAGACTCAGGGT 541
 QY 1380 CATGTATGTCCTTGGGGTTGAATTTCAACATGCTGAGTTTATGTGTTCTTGGC 1439
 DB 542 CATGTATGTCCTTGGGGTTGAATTTCAACATGCTGAGTTTATGTGTTCTTGGC 601
 QY 1440 TGTATGCACAGATCAAGGAAAAAGCCGGGCTTGGCGTATCAAGATCTCAATGGCC 1499
 DB 602 TGTATGCACAGATCAAGGAAAAAGCCGGGCTTGGCGTATCAAGATCTCAATGGCC 661
 QY 1500 ACATTTACCAAGATCAACTGCAATTTGATGCGGATGCGAGTGAAGAGTGAAGCAATTC 1559
 DB 662 ACATTTACCAAGATCAACTGCAATTTGATGCGGATGCGAGTGAAGAGTGAAGCAATTC 721
 QY 1560 CAGCGCTCAGTTCCATATCAATCAAGATTAAAACTGCAAGATTGGAACTTGGG 1619
 DB 722 CAGCGCTCAGTTCCATATCAATCAAGATTAAAACTGCAAGATTGGAACTTGGG 781
 QY 1620 TCACCTTGAATGTTTGAAGTCAACCGGATATCACTTCCAGATCTCAATCAATCCGT 1679
 DB 782 TCACCTTGAATGTTTGAAGTCAACCGGATATCACTTCCAGATCTCAATCAATCCGT 841
 QY 1680 TTTTCAGTCTAA 1690
 DB 842 TTTTCAGTCTAA 852
 RESULT 12
 AX009482 838 bp DNA linear PAT 06-SEP-2000
 LOCUS Sequence 2 from Patent WO961634.
 DEFINITION AX009482
 VERSION AX009482.1 GI:9996767
 KEYWORDS
 SOURCE Vibrio cholerae
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio.
 REFERENCE
 AUTHORS Carlin, N. and Lebens, M.R.
 TITLE Method of producing thy a<->strains of vibrio cholerae, such strains and their use
 JOURNAL Patent: WO 961634-A 2 02-DEC-1999;
 CARLIN NILS (SE); SBL VACCIN AB (SE); LEBENS MICHAEL R (SE)
 FEATURES
 source 1. 838
 /organism="Vibrio cholerae"
 /mol_type="unassigned DNA"
 /db_xref="taxon:666"
 ORIGIN
 Query Match 28.8%; Score 838; DB 6; Length 838;
 Best Local Similarity 100.0%; Pred. No. 1.9e-248; Indels 0; Gaps 0;
 Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAAGTTTGTATGCTCAGGGTATCTGAGTTTCCCAATATGACCCCGTATTTGT 60
 DB 1 GAGAAAGTTTGTATGCTCAGGGTATCTGAGTTTCCCAATATGACCCCGTATTTGT 60
 QY 61 TTGCATCGCCCTCTAGCGGTGGCTGTGATGCTTGAATTTGGTGGTTTCTTTT 120
 DB 61 TTGCATCGCCCTCTAGCGGTGGCTGTGATGCTTGAATTTGGTGGTTTCTTTT 120
 QY 61 TTGCATCGCCCTCTAGCGGTGGCTGTGATGCTTGAATTTGGTGGTTTCTTTT 120
 DB 61 TTGCATCGCCCTCTAGCGGTGGCTGTGATGCTTGAATTTGGTGGTTTCTTTT 120
 QY 121 TGTATGATGTTGGCAATCGCGAGCGATCGCGGGCAGTGGTTGAGCGGTGAGCA 180
 DB 121 TGTATGATGTTGGCAATCGCGAGCGATCGCGGGCAGTGGTTGAGCGGTGAGCA 180
 QY 181 AGTCTCTGACTTGTATTCGCGGGCTTTTAAAGTGTAGTATCGATCGGCGAGTTGTTA 240
 DB 181 AGTCTCTGACTTGTATTCGCGGGCTTTTAAAGTGTAGTATCGGCGAGTTGTTA 240

QY 241 TGTATGATCTTCAATTTTATGATCTGTTCTTGTGACCTCTTATTTATCAAGTGTG 300
 DB 241 TGTATGATCTTCAATTTTATGATCTGTTCTTGTGACCTCTTATTTATCAAGTGTG 300
 QY 301 GACTGCGGCATGCTCTTCCACGCGGCTTATTTGGTGTATCACCGCATGTTCTGTA 360
 DB 301 GACTGCGGCATGCTCTTCCACGCGGCTTATTTGGTGTATCACCGCATGTTCTGTA 360
 QY 361 TGGCGGTAAAAACAACGACCTTTTGTGTGGCCCATTTTGTGGCCCTTTAGTGGC 420
 DB 361 TGGCGGTAAAAACAACGACCTTTTGTGTGGCCCATTTTGTGGCCCTTTAGTGGC 420
 QY 421 ATTGCGTTGGGGATGGGACGTATGCTGATCTTATGAAATAGTGAATTTGGGGAGAGT 480
 DB 421 ATTGCGTTGGGGATGGGACGTATGCTGATCTTATGAAATAGTGAATTTGGGGAGAGT 480
 QY 481 AAGGATGCTGCTTGGGCTTTTGTATTCCTATATGATGGGCCCATGCGGCATCTTC 540
 DB 481 AAGGATGCTGCTTGGGCTTTTGTATTCCTATATGATGGGCCCATGCGGCATCTTC 540
 QY 541 AAGGATGCTGCTTGGGCTTTTGTATTCCTATATGATGGGCCCATGCGGCATCTTC 600
 DB 541 AAGGATGCTGCTTGGGCTTTTGTATTCCTATATGATGGGCCCATGCGGCATCTTC 600
 QY 601 TGTAAACCTGCTCGCTAGGAGGAGGTATCGGACGTGTTTATGCTGATAGGATATT 660
 DB 601 TGTAAACCTGCTCGCTAGGAGGAGGTATCGGACGTGTTTATGCTGATAGGATATT 660
 QY 661 CCGCTTCTTGTGAAATAGTCCGTAGGACAGATCTCAGTTGGTCTGTTGGTGGCT 720
 DB 661 CCGCTTCTTGTGAAATAGTCCGTAGGACAGATCTCAGTTGGTCTGTTGGTGGCT 720
 QY 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATGATCATGCGTATTTGATGATGCT 780
 DB 721 CATTTCAATGAGGCAATCTCTCTTACCTATGATGATCATGCGTATTTGATGATGCT 780
 QY 781 TTTGCTTACAGAGCGGCTTGTATCAAGACCGGTAGAGCAAAATGGGTAGTAG 838
 DB 781 TTTGCTTACAGAGCGGCTTGTATCAAGACCGGTAGAGCAAAATGGGTAGTAG 838
 RESULT 13
 AE006043/c 10719 bp DNA linear BCT 08-MAR-2001
 LOCUS AE006043 AE004439
 DEFINITION Pasteurella multocida PM70 section 10 of 204 of the complete genome.
 ACCESSION AE006043.1 GI:12720282
 VERSION AE006043.1 GI:12720282
 KEYWORDS
 SOURCE Pasteurella multocida
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Pasteurella.
 REFERENCE
 AUTHORS May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whitlam, T.S. and Kapur, V.
 TITLE Complete genomic sequence of Pasteurella multocida, PM70
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
 MEDLINE 11248100
 PUBMED 11248100
 REFERENCE
 AUTHORS Zhang, Q. and Kapur, V.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA
 FEATURES
 source 1. 10719
 /organism="Pasteurella multocida"
 /mol_type="genomic DNA"
 /strain="PM70"
 /db_xref="taxon:747"

gene 225. .1076
/gene="PM0074"
CDS 225. .1076
/gene="PM0074"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02158.1"
/db_xref="GI:12720283"
/translation="MKSSESTCWSPTRKAOIGEDVGVYKATKKOIYSPESAI
PAGAFIATAPVYATOTQNGNDAPMGGLTKYGVIFSGVIMVVCSELFTSSMTAR
VSGRISTIQMLKNMLVIGGNVGLFVLLVTRAGQIMANGQWGLTILTKTAQKH
HTWTEAFALGIFCNITMVCIAVMTYAGTKLLDKAVITMLPISMFVASGPEHSVAMFM
IPMGMMIAHFASPEFWQAVSLNPAQFADLDWYHFVKNLIPVLLGNIVGALCVALTQ
WYINRP"
gene 1156. .3486
/gene="pf1b"
/note="synonym: PM0075"
CDS 1156. .3486
/gene="pf1b"
/codon_start=1
/transl_table=11
/product="pf1b"
/protein_id="AAK02159.1"
/db_xref="GI:12720284"
/translation="MMSQMLNETQOKANEGETGDMQTEVNVVRDPIQRNVTPEYEGDES
FLADYEAATTKLMNDVMEKIKYENTHREYDIDCPTSTITSHAGYIDKSLKIVGL
OTDAPKRAIMPRGGIINMYKGSCKYRREIKREVEQIPTEYRKTNGOVDPDYITVL
RCRSGVITGLPDAYGGRGITDYRYRMALYGDPIMKRKNQFTSLQKLRGEDIQA
TTLQREIABQHRALGKMKEMASGYDISEPATNAHEVQTTYPATLAAYSGNGAA
MSFGVSTFLDIYIERDLKAKRITEOEQELIDHLVMKLRMVRFLRTEBYQOLFESGD
MWATEFLAGMGIDGRITLVTKNSFRILHTLYTWGPSPPEPMLITLMSEKLPBEGKRYAK
VSIDTSVQYENDDLMRPDFONDYALACVSPMIVGKMOPFGARAMLAKILLAYIN
GVYDSESGDOVGPKTDPIITSEYLDYDWDWTRTLDSPMDLAKOYVTLAINIHHMHKVA
YEAALMALHDRVPRPTMACGIAGLSVAADSLSAITYAKYKPRGDIETIKKRAGVYGI
AADVVAIDETISEYIPQFGKNDNRVDEIACDLVERWKCIQKLGITRNATPTOSVLTIT
SNVVYKGTGNTPDGRSGAPPGPANPMHGDQGAVALTSVAKLPFAAKADISY
TFSIVPNAIGKDYEAQKNLIGLMDGYFHEHTITEGQHLNVNVRNREMLDAMENPE
KYPOLIRVSGYAVAFNSLTKEQQQDVITRTPTQAM"
gene 3659. .5698
/gene="est"
/note="synonym: PM0076"
CDS 3659. .5698
/gene="est"
/codon_start=1
/transl_table=11
/product="est"
/protein_id="AAK02160.1"
/db_xref="GI:12720285"
/translation="MKINTLSFALPCSLSSAALAKDVVFGDSLSDMQOTGNKKASY
LKADGSYHLYDEYLAQAGKGLIPSTQGLNAVYSGGVIYGAHNTRTAEOPHLLEK
QINELYIAHPVKEEALHILMAGNDLATVATATVKTTPBEKQAVLASINTMAQMAO
QNGALQAGVNOITAPTTIPNVYTPPEFPDKLGBAAGAOIKOASYGLIKOSDPVANPKA
AAEQLLVQPTKNLEEFERKRVQTLERKAEDFYNSRYWFTQIALSSAGINAKIATLTI
KEYKTIVBAPQAATAFANASITSALNOVGNIIVRIDTGLMDMMTRPAEYGLINTTT
VYCKSTIADPAPACKBEDQTAQSRLTFDSEPHPTAHKAMSQDILNVLQPKMGI
LTOIVQOQTELDLPIRTESNRHRLQOQASQAGVITLAAOKGSHLSHVGAQVFN
QMOIALIVSOKODYQOGLINTNONTNUTSTLRVDADMMWISGLVONNTLLTIDRY
AATGSHAQQSARETAESFSKGLPAGYEKFKALGIALIADLNKTKTVAGAGEONRG
ITQKQFAARTKSVSGVGFDIRYQASWQPTVTRVWVENNDIPTIOAGINSFSM
TVLPBODNOMVNI VAGLHPKPISSGFHANLARRDLGRKNALNTTLLQAGIGFTP"
gene 5862. .6602
/gene="act"
/note="synonym: PM0077"
CDS 5862. .6602
/gene="act"
/codon_start=1
/transl_table=11
/product="act"
/protein_id="AAK02161.1"
/db_xref="GI:12720286"
/translation="MAVVGRIHSYSSCGTVDPGIRFILFMQGLMRQCYCHNRDWD
LHAGKEITVEELMKEVVYTRHFMNASGGGVITASGSAVLQAEFVADWPSACKKEGHT

gene CULDNCFVRNYDVHIDELLDVTDLVLIDLKQJNDKIHQNLICVFNKRTLEPQYIAYR
NQPYWIRYVAVPGYTDADEDIHLGHFIKMKNIKEVELLYRHLGAKHWEAMEKYE
LEEYVPTKESLEBHKISLESYGHITIK"
CDS complement (6819. .7319)
/gene="PM0078"
complement (6819. .7319)
/gene="PM0078"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02162.1"
/db_xref="GI:12720287"
/translation="MQSDMDKMKRRLRLADKABQJGEIPVGAUVYDEQANLLBEGH
NLSTSESDPTNAHAIVALROGQRLQYRLLNTTLYTLEPTWCAGALHSRIKRLV
FGASDYKTGAHSAFRHFDDYKONHALEITAGVQEBESOKISAFQKRREQOKIARQ
KSHLET"
gene complement (7334. .8185)
/gene="thya"
/note="synonym: PM0079"
CDS complement (7334. .8185)
/gene="thya"
/codon_start=1
/transl_table=11
/product="thya"
/protein_id="AAK02163.1"
/db_xref="GI:12720288"
/translation="IMKHYLELCQRIVDEGVWENARTGKCLTVIDADLTYYDVAANO
PFLITRRSYWAAIAEFLGYIRGYDNADEFKLGTKTMDAANANSAMLNPHKRGTD
DMGRVYVGOGAAMRKPNGETVDOLRKIVNLRNGLDRGEIMTEFNGEPELGLCRPC
MHTYTPSLSGTLYLTYSQBSCDPILNENQIOVPTFLMAQDITCKKGAAYHKII
NAHYEDDLMKVQLRRDFPLPOLHINPEIKTLEDLETWVTMDPFTYGYOCHDA
IKYPSF"
gene complement (8182. .8994)
/gene="tgt"
/note="synonym: PM0080"
CDS complement (8182. .8994)
/gene="tgt"
/codon_start=1
/transl_table=11
/product="tgt"
/protein_id="AAK02164.1"
/db_xref="GI:12720289"
/translation="MQSYRSPQFPQDPVYIFETGPIGLRWYGLMYLGLFLPARNLAVR
ANQTSQWTTDQVDSLFPNGFWGVFLQRIQYVLYQPDFLQDPAYLFRWEGMSF
HGGILGIVTAMLTIRAKIQKRGFMQTAQFVALLPFGQMGRIQNF INDELMRGVTDVP
WAVLPSSGCVLRPHSOLYEAVLEGIYVFLIANYIKKPRPIGATAGLFLIGYGIFFR
IVERPREBDALQGLYFGQHISMGOILSTPMILIGAVIMLVAYQSAKKREIL"
gene complement (9004. .9798)
/gene="PM0081"
CDS complement (9004. .9798)
/gene="PM0081"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02165.1"
/db_xref="GI:12720290"
/translation="MFTFFELICLAVNGAFGLAGLFGIGGLVTPVLVYLMPKXVGV
PEAWMACALTSFATVITGFSASORHKKGVNMDVKYLAIVIMTTFISGPFIS
YLPKDISARLACIVTALIKMPLSVKRPYTKALTKRSTIIAGILGMASAGYIGG
GGFIVPPLANSGITDMKKAIGSSAYCGSLIAIAGNLSPNIAAMKYBGLPEYMSGYVLP
ALVGITLTSITSKIGATVISILPVPILKCAFALMLTILAINMLIK"
gene complement (9801. .10397)
/gene="PM0082"
complement (9801. .10397)
/gene="PM0082"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK02166.1"
/db_xref="GI:12720291"
/translation="MIDPDFRPNVGVIVICNSKQOVLAKRYGQNSMQFROGGINDNE
SABQAMRLELEBYGLSKDVKILYISGMLYKLPKRLIATYDSKPVCITGQKORFLL
QVDEKXINNQSSKSPFDEGMWVSFWYPRQYVVSFKDEVYRAMEFASVLPDGAK

ORIGIN

ENLSSKSNESDKTHHTTKSTFLTKSHKPHKSRG"

Query Match 28.0%; Score 814.8; DB 1; Length 10719;

Best Local Similarity 68.7%; Pred. No. 5.2e-241;

Matches 1161; Conservative 0; Mismatches 507; Indels 22; Gaps 2;

QY 4 AAGGTGTGATGCTCAGGGTATCTGACATTTCCCAATATTTGACCCCGATTTGTTTC 63
 DB 9004 AAGGAATTTATGCAATGCAATTTATTTTCAATTTCCCAATTTGATCCGTCATTTTGA 8945
 QY 64 GATCGGCCCTTACGCGTGGCTGTATATGCTTATTTGATTTGTTTCCCTTTTTC 123
 DB 8944 AATCGGCTCATGCTTACGTTGATGCTTATATGATCTGCTGGGTTTCTCTTTC 8885
 QY 124 TATGTGTTGGCCATGCGGACCGATCGCGGGGAGTGTGTGACCGCTGACAGT 183
 DB 8884 CCGTTGCTTGGCGGTGAAAGTGTCTATCAACAGGACGTGGCTGACACAGCCAGT 8825
 QY 184 CTCGACTTGTATTTGCGCGGCTTTTATGAGTATGATCGGTGGCCGATTTGATAT 243
 DB 8824 GATAGCTCTTATTTTACGTTTATGCGCTTTTATGATGACGATTTGTTTACGT 8765
 QY 244 GATCTTCTACATTTGATCTGTCTTGTGACCTCTTATTTATTTCAAGTGTGAC 303
 DB 8764 GTTGTTTTATCAATTTGATTTATTTCTTACAAAGATCGGCTATTTATTCGTTTGGGA 8705
 QY 304 TGGCGGATGCTCTTCCACGCGGCTTATTTGGTGTGATACCGCCATTTCTGTATGC 363
 DB 8704 AGCGGATGTCTTTCATGTGTGCTTAAATTTGTGTGATTTTGTGCAATTTATTCACAGC 8645
 QY 364 GCGTAAACCAACGACCTTCTTTTGTGTGTGCGATTTTGTGCGCCCTTATGTCATT 423
 DB 8644 TAAATATCAAAACGCGGTTTCTGCAACGCGGATTTTGTGCGCCCTTATTCATT 8585
 QY 424 CCGTTTGGGATGGAAGATTCGCTATCTTATGAATGAACTTTGGGAGACGATGAC 483
 DB 8584 TGGCTTAGGATGGAAGATTCGCTATCTTATGAATGAACTTTGGGAGGCTGAC 8525
 QY 484 GATGTGCTTGGGCTTTTGTATTTCTTATTTGTTGCGCCATCGCGGCTATCTTCA 543
 DB 8524 GATGTGCTTGGGCTTTTGTATTTCTTATTTGTTGCGCCATCGCGGCTATCTTCA 8465
 QY 544 GCTTTATGATTTGCTTATGAGCGGCTTCTGTTCTTATTTCTTATTTGTTTATG 603
 DB 8464 GCTATGAGGAGCGGATTTGAGGCGATTTCTTATTTCTTATTTGTTTATGATTCAA 8405
 QY 604 TAAACCTGCTCCGCTAGGACGCTATTCGAGCTGTTTATGCTGATACGTTACATTCG 663
 DB 8404 AAGGCTCGACCAATGCGGCGCAACCGTGTATTTCTTATCTGCTTACGTTATTTCCG 8345
 QY 664 CTTCCTTGTGAATGATCGCTGAGCCAGTGTCAATGTTGTTCT--GTTTGTGGCTT 720
 DB 8344 TTTTATGCTTGAATTTTTCGTAGCGCTGATGCCCAATTTGCTTATTTTGTGACACA 8285
 QY 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATGATGATTTTATGATGCT 780
 DB 8284 TATCTCATGAGGCAATTTCTCTCACCAATGATTTCTATTTGTCGGTATTCATGTT 8225
 QY 781 TTGCTTTTCAACGCGGCTTGTATCAAGACCGTGTAGCAAAATAGGTTAGTTAGT 840
 DB 8224 AATGCGATATCAAAATG-----CGGTAATAAAAGGAATATTTAT 8184
 QY 841 GAAACGATTTATGATCTTTGTACGCGCATGCTGATCAAGTGTGTTGTTGTTGAAATGA 900
 DB 8183 GAAACGATTTATGAACTTTTGCACAGTATGCTTATGAAAGTGTGTTGTTGTTGAAATGC 8124
 QY 901 AGGAGGAGGAGGATTTGATGCTGATTAATGCGATTTATGCTTATGATGATGAGGCA 960
 DB 8123 ACCGATGAGGAAAGTGTGCTGACCGGATTTGAGGATCTCACTCACTGATGTGCA 8064
 QY 961 CAATCAATTTCTCTGATGATCAAGCAAGATTTTGTGAAAGTCTCCGTACCGCATTT 1020

DB 8063 TAAACAATTTCCGTTGATCAACACAGCAAAAGCTATTGAAAACCGGCAATTCAGAAAT 8004
 QY 1021 GCTCGGCTATTTTCGTTGATGATATGCGGCGGATTTTTCGCAATTTGATGACAAAC 1080
 DB 8003 TTTAGGTTATATCCGTTGATGATGATGCGGAGATTTTCGAAACTGCGACCAAAAC 7944
 QY 1081 CTGGATGCTAATGCCAATTTAAACCAAGATGCTCAACATCTTACCTGTTAAAGTGA 1140
 DB 7943 CTGGATGCGCAATGCTATGAAACAGTGCATGCTCAATATCCGACCGCAAAAGCAC 7884
 QY 1141 GATYACATGAGACGCGTATGCTGTGATGCTTACGCTTACGCTTACGCTTACGCT 1200
 DB 7883 CGATGATATGAGGACGTTTATGCTTACGCTTACGCTTACGCTTACGCTTACGCT 7824
 QY 1201 TCATATTTACCGTTGAAAAAGTTGATGATTTGAGCCGCGGCTTATGATGACGAG 1260
 DB 7823 AACCGTCAATTTGCGCAAAATCTCAATATTTGCGCAACGCGATGATGATGATGAT 7764
 QY 1261 TGAATTTCTTAACTTTTACAAATCCGCTGATATTTCAATGCGGCTTTCGCTTGCAT 1320
 DB 7763 TGAATATGATGATCTTTTCAACCGGCTGATATTTGAACTGCGTTGCTTACGCTTGCAT 7704
 QY 1321 GTACAGCCATCATTTTCAATGCTGCGGATATCTTATCTCAACAGTATCAAGCTTTC 1380
 DB 7703 GCACACCAACACTTCTCAATCTCGTATACCTCTATCTCAACAGTATCAAGCTTTC 7644
 QY 1381 ATGTATGTCCTTGGGCTTGAATTTCAACATGTCAGGTTATGTTCTTCTGCTT 1440
 DB 7643 TGTGATGTCGCTTGGCTTGAATTTCAACAAATTTCAAGTGTCACTTTCTTCTT 7584
 QY 1441 GATGCAACATCAACAGGAAAAAGCCGCTTGGCTTACCAAGATGCTCAATGCGCA 1500
 DB 7583 GATGCAACATTTACGCGCAAAAGCGGTAAAGGCTTACCAAAATCATCAATGACACA 7524
 QY 1501 CATTTACCAAGATCAATGCTGAATTTGCGGATGTCAGCTTAAAGTGAAGCATTTCC 1560
 DB 7523 CATTTATGAAGATCAATGCTGATCTCATGAAAAATGCAATTTAAAGTGAAGCATTTCC 7464
 QY 1561 AGGCTCTGCTTCAATGATCAATCCAAAGATTTAAACATGCGAGATTTGAAACCTTGG 1620
 DB 7463 GTTGCACAAATTAACATTTACCCGAGATTTAAACCTGCAAGATTTGAAACCTTGG 7404
 QY 1621 CACTTGAATGATTTGATGCTACCGGATATGATGCTTCAAGATCTTCAATACCGCT 1680
 DB 7403 GACTATGATGATTTTATGATGCTACCGGCTTATCAATGCTCAATGATGATTTCTT 7344
 QY 1681 TTCAGTCTAA 1690
 DB 7343 CTCGTGTAA 7334

RESULT 14
 AF064791 1893 bp DNA linear BCT 31-DEC-2001

LOCUS AF064791
 DEFINITION Pasteurella multocida thymidylate synthase (thyA) gene, complete cds.

ACCESSION AF064791
 VERSION AF064791.1 GI:18000047

KEYWORDS Pasteurella multocida
 SOURCE Pasteurella multocida
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Pasteurella.

REFERENCE 1 (bases 1 to 1893)
 AUTHORS Moreno, J.A., Bosch, M., Badiola, I., Llagostera, M. and Barbe, J.

TITLE Direct Submission
 JOURNAL Submitted (12-MAY-1998) Departamento de Genética y Microbiología (Unidad de Microbiología), Universidad Autónoma de Barcelona, Edificio Cn, Bellaterra, Barcelona 08193, Spain

FEATURES
 source
 1..1893
 /organism="Pasteurella multocida"
 /mol_type="genomic DNA"

gene
CDS

```

/strain="PM25"
/specific host="bovine"
/db xref="Taxon:747"
826..1677
/gene="hlyA"
826..1677
/gene="hlyA"
/EC number="2.1.1.45"
/codon start=1
/transl table=1
/product="chymotrypsin-like synthase"
/protein_id="AAB54880.1"
/db xref="GI:18000048"
/translation="MKHYLELCORIVDEGVWENRTEKRLIVIDADLYDVANNQF
PLITRKSVMKAAIAEPLGYIRGYDNADPRKLGKTVMDANENSAMLANPBRKGTD
DMGRVYGVGGARKENGEIVDOLKIVINILNNGIDBGEIMTPNPEFELGRLPC
MHTTFPSLIGTLVLTYSORSCDVPICAFNPNOIOVFTPLATMAOI TEKKAKAYKII
NAHIYEDOLDIMKRVQLKREPPPLFQLHINFEIKLELDELTVMTWDDPKVTGYQCHDA
IKYPSV"

```

ORIGIN

Query Match 26.6%; Score 775.2; DB 1; Length 1893;
 Best Local Similarity 67.9%; Pred. No. 7.8e-229;
 Matches 1119; Conservative 2; Mismatches 505; Indels 22; Gaps 2;

```

46 TGACCCCGATATGTTTCGATCGCCCTCTAGCGGTGCGTGTATGCTGATGATTT 105
49 TGATCCCGATATTTTGAATCGGTCTATCGTTTACGTTGTATGCTTATGATCT 108
106 GGTGGGTTTCTTTTGTCTATGTGTTGGCCATCGCGAGGATCGCGGGGAGTGG 165
109 GCTGGGGTTTCTTTGCGCCGTGGCTTGCGGTGAACGCTGTAACAAACAGGAGTGG 168
166 TTGACGGGTGAGCAAGTCTGACTGTGTTATTCGCGGCTTTTAAAGTATGATCGG 225
169 CTGACAAACAGCAAGTGAGTACCTGTATTTAAAGGTTTAAAGCGTTTATTTAGG 228
226 TGCGCGAGTTGGTTATGTGATCTTCTACAAATTTGATCTGTCTGTCGACCTCTT 285
229 TGCGGCGATTTGGTTATGTTGTTTATTCATTTCAATTTCTTAAAGATCTGCGCT 288
286 TTTTATCAAGTGTGAGCTGCGGCGATGTCCTTCAACGCGGCTTATTTGGTGTATCA 345
289 TTTTATTCGTTGTTGGGAAGGCGGTATGCTTTCATGCGGCTTAAATTTGGTGTATTT 348
346 GCGCATGTTCTGTATGCGCGCTTAAACCAACGACCTTCTTGGTGGCGGATTTTGT 405
349 AGCCATGTTATCAAGGCTTAAATTAACAAACGCGGCTTCTGCAACGCGGATTTTGT 408
406 TGCCCTTTAGTGCATTCGTTTGGGATVGGAGGTATCGGTAACTTTATGATATGTA 465
409 GGCACGTTGATTCATTTGGCTTAGATATGGAGATCGCACTTTATCAATGACGA 468
466 ACTTTGGGAGCAGATGAGATGTCCTTTGGGCTTTTGTATTCCTAATGTGGCCACT 525
469 ATTGTGGGGCGTGTACGAGATGTCATGGCGACCTTATTCATCATGAGGGGTATTT 528
526 GCCCGGCATCTTACAGCTTATGAAATTCGCTTAAAGGCGGTATTCGTTCTTAT 585
529 ACCACGCAATCCMTACAGCTATACGAGGCGTATTAAGAGCATTTGTCATTTCTTAT 588
586 TCTTAATTTGTTATTTGTAACCTGTCGCTGACGAGCGATCCGATCTGTTTATGC 645
589 CCYGAATTTGTAATCAAAAACCTCGACCAATCGGTGCAACSTTGGTTATTTCTACT 648
646 TGGATACGGTACATTCGCTTCTGTTGGAATACGTCGTTGACCAATGCTCAATGGG 705
649 CGGTACCGTATTTTTCGTTTATCGTTGAGTTTTCGTTGAGCTGATGCCCAATTAGG 708
706 TCT---GTTTGGTGCCTCATTTCAATGGGGCAATTCCTCTTACCATAGGTATCAT 762
709 CTTAATATTTGGTCAACATCTCTATGAGGCAAAATCTCTCACACCAATGATTTCTAT 768

```

```

QY 763 CGTATTTTGTATGATGTTTGTCTTACACGCGGCTTTGTATCAACGCTGTAGCAC 822
DB 769 TGTGTGCGGTATTCATGTTAGTGCGCATATCAAAAGT-----CGGGTA 809
QY 823 AAAATAGGTAGTTAGTGGAACAGTATTTTGTATCTTTGTCAGGCACTGTCATCAAG 882
DB 810 AAAAAGGTAAATTTTATGAAACCTATTTAGAACTTTTGGCAACGTATCGTTGATGAAG 869
QY 883 TGTTTGGTTGAAAATGAAACGAGCGCAACGTTTGTGATCTGATTAATGCGCATTT 942
DB 870 TGTTTGGTTGAAAATGAAACGAGTGAACGTTTGTGACGCTGATTAATGACGATCT 929
QY 943 GACCTACGATGTGGCAACATCACTTTCTTGTAGTACCTACGCAAGACTTTTGGAA 1002
DB 930 CACCTACGATGTGGCAACATCACTTTCTTGTAGTACCTACGCAAGACTTTTGGAA 989
QY 1003 AGCTGCGGTACCGAGTGTGCTGCGGTATATGTTGTTAGATATGCGGGGATTTTGG 1062
DB 990 AGCGCAATTCAGAAATTTTATGTTATATCGTGTATATGCAATGCGGAGATTTCCG 1049
QY 1063 CCAATTTAGTACCAAAACCTGAGATGCTAATGCAATTTAAACCAAGCATGCTCAACAA 1122
DB 1050 CAATCTGCGCACCAAAACCTGAGATGCGCAATGTAATGAAACAGTGCATGCTCAATTA 1109
QY 1123 TCTTTACGTTAAGGTGAGATGACATGGAGCGGTATGTGTTTCAAGGTAGACTTG 1182
DB 1110 TCCGCAACGCAAAAGGCAACCGATGATATGGAGCTGTTATGCGGTACAAAGGCGTGCATG 1169
QY 1183 GCTTAAGCTATAGTGTGCTATTTGACCACTTGAATAAAGTTGTATGATTTGAGCGG 1242
DB 1170 GCGCAACCGAAGCGGTGAACCGTGCATCAATGCGCAAAATCGTCAATTAATCTGCGCAA 1229
QY 1243 TGCGGTGATGACGAGGTGAATTTCTTAACTTTTACATTCGAGGTGAATTTTCAATGG 1302
DB 1230 CGGATTTGATGACGTTGTTGATATGATGATGATGATGATGATGATGATGATGATGATG 1289
QY 1303 GTTGTGCGCTTGCATGATGACAGCATATTTTCACTGCTGCGGAGTATCTTGTATCT 1362
DB 1290 TTTGCTTACGCTTGCATGATGACAGCATATTTTCACTGCTGCGGAGTATCTTGTATCT 1349
QY 1363 CAACAGTATCAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1422
DB 1350 CACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
QY 1423 TTTATGTTTCTTGTGCTGATGAGCAAGATGACAGGAAAAAGCGGCTTGGGTATCA 1482
DB 1410 GTTACACTTTTGTGCTTGTATGAGCAAAATTAACGCGGAAAAAGCGGTTAAGCGGTACCA 1469
QY 1483 CAAGATGCTATGAGGACATTTTACCAAGATGCACTGATTTGATGCGCATGTGCACT 1542
DB 1470 CAATATCATCATGATGACATTTTATGAAGATCACTGATCTCATGAAAAATGTGCAAT 1529
QY 1543 AAAACGTAGGCAATTCAGAGGCGCTGCTTCAATATCAATGCAAAATTAACACTGCA 1602
DB 1530 AAAACGTAGGCAATTCAGAGGCGCTGCTTCAATATCAATGCAAAATTAACACTGCA 1589
QY 1603 GGAATTTGAAAATCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1662
DB 1590 AGATTTGAAAATCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1649
QY 1663 TCTTATTCATATCCCGTTTTCAGTCTAA 1690
DB 1650 TGCCATTAATATCTCTTCTCTGTGTAA 1677

```

RESULT 15
 U32772 11545 bp DNA linear BCT 29-MAY-1998
 LOCUS Haemophilus influenzae Rd section 87 of 163 of the complete genome.
 DEFINITION Haemophilus influenzae Rd section 87 of 163 of the complete genome.
 ACCSSION U32772.142023
 VERSION U32772.1 GI:1573918
 KEYWORDS
 SOURCE Haemophilus influenzae Rd KW20

ORGANISM Haemophilus influenzae Rd KW20
Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus.

REFERENCE 1 (bases 1 to 11545)
AUTHORS Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirschnes, E.F., Kierlavage, A.R., Bult, C.J., Tomb, J., Dougherty, B.A., Gerick, J.M., McKenney, K., Sutton, G.G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shiley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T., Hanna, M.C., Nguyen, D.T., Sander, D.M., Bratton, R.C., Fine, L.D., Fritchman, J.L., Fummann, J.L., Geoghagen, N.S., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.
TITLE Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL Science 269 (5223), 496-512 (1995)
MEDLINE 95350630
PUBMED 7542800

REFERENCE 2 (bases 1 to 11545)
AUTHORS Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Bordovsky, M., Rudd, K.E. and Koonin, E.V.
TITLE Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
JOURNAL Curr. Biol. 6 (3), 279-291 (1996)
PUBMED 8805245

REFERENCE 3 (bases 1 to 11545)
AUTHORS White, O., Clayton, R.A., Kierlavage, A.R. and Fleischmann, R.D.
TITLE Direct Submision
JOURNAL Submitted (25-JUL-1995) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
PUBMED 4 (bases 1 to 11545)

REFERENCE 4 (bases 1 to 11545)
AUTHORS White, O., Clayton, R.A., Kierlavage, A.R. and Fleischmann, R.D.
TITLE Direct Submision
JOURNAL Submitted (27-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
PUBMED

REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
5 (bases 1 to 11545)

REFERENCE 5 (bases 1 to 11545)
AUTHORS White, O., Clayton, R.A., Kierlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
TITLE Direct Submision
JOURNAL Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
PUBMED

REMARK The whole genome was shifted by 588 nucleotides for a new start On Sep 30, 1996 this sequence version replaced gi:1221622.
COMMENT FEATURES
source location/Qualifiers
1..11545
/organism="Haemophilus influenzae Rd KW20"
/mol_type="genomic DNA"
/db_xref="taxon:71421"
complement(122..604)
/gene="HI0899"
complement(122..604)
/gene="HI0899"
/note="similar to GB:D10483 SP:P00379 GB:J01609 GB:V00276 GB:X05108 percent identity: 53.16; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="dihydrofolate reductase (folA)"
/protein_id="AAC22559.1"
/db_xref="GI:157319"
/translation="MTFSLIVATTLANNVIGKONPWHLPADLANFRONTTCKPVMGRTFESIGRLPRTNIVISRLPFEEGVWDSFESAVNFRDDEIMLIGGELFPKQYLPRADKLYLTQITELDDTFPQLNMEWEIEPDEYKADQNRYDCRPLILTRK"

gene 705..1811
/gene="HI0900"

CDS 705..1811
/note="HI0900"
/note="similar to GB:U00096 SP:P07005 PID:1208988 PID:1552810 PID:1786437 percent identity: 65.57; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="glutamate 5-kinase (gamma-glutamyl kinase) (proB)"
/protein_id="AAC22560.1"
/db_xref="GI:1573920"
/translation="WNKKTIVKFGTSTLTQSPPLNSPHMIEIVRQIAQLHNDGFRIVISGAAAGRHLYLNHPOLPTTIASKQILLAVGQSOLQIAWEKFLAYDHIQGLLTRAIDEDBERFLNADDTLYALDDNIIPVINDAVVATAEIVGNDNISALVALIVQABOYLTDQGLPDSDRKNPEAKLIPVEQITDHSIAGSGTNGTGMWKTI AADVATRSIGETITIAPKRPNVIALALAEQNGITFIAHQSDRLSRRCMTLPASAG IITIDNGAONALIEONKSLPAGIITNIGRSRSGVVKIRTQSGKDIADGMPRTNSDA LQILIGKRSADIENTVLGYEYGAVMHRDMILTS"

gene 1892..2482
/gene="HI0901"
/note="HI0901"
/note="similar to PID:882723 GB:U00096 SP:Q46930 PID:1789194 percent identity: 77.71; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22561.1"
/db_xref="GI:1573921"
/translation="MIDFGYRPVNGIVYICNRKQGVLAKEGQNSQWOPQGINDE SAEOAMREIHEHVEGLQPKDRLVLYGVSIKRYKLPRKLIRDSKPGICGKQWFL QLVSDKRNINQTTSPSPEDGRWVSFPYPRQVVSFPKDYRKWKESASLITFDNP LIFSASREANSLSHSANKKTSOTKTKHFKSKQ"

gene 2482..3276
/gene="HI0902"
/note="HI0902"
/note="similar to GB:L77117 SP:O57883 PID:1591145 percent identity: 28.84; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC22562.1"
/db_xref="GI:1573922"
/translation="MTFPIILCLVGLAGLAGLFGIGGLVIVPTLVLYLPIDVP ESLMSTALGTSFATVITGIGSAQRHKLGNIVQAVRIAPVIMLSVFCGLFGR LDRBISAKIPACLVVYLTAKWVLSIKQOVTTKSLPLPSVIGILIGMSAAGIG GEFIVPLTAGGINIKOAISSAPCGMILISGMSFVIGSGNLMREYSGLIYLP AVLGITATSFFTSKIGASATKALPVSTLKGFPALFLIVVAINMFLK"

gene 3285..4091
/gene="HI0904"
/note="HI0904"
/note="similar to SP:P37149 GB:U12289 PID:516036 PID:882720 GB:U00096 percent identity: 62.78; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="prolipoprotein diacylglycerol transferase (lgt)"
/protein_id="AAC22563.1"
/db_xref="GI:1573923"
/translation="WNSNYLLPLPHDP81FTLGDNSNIGLRMYGLMYLLGPFARMLAV BRANPNNGMTVDQVDSLLPFGMGVPIFGHGVDFVYNLHFLQELPYLFLPMVGEMN SFHGLIGVIVAMITYSOSKRNFWQCTFAPALPLPFGIGRINPILNLEMGRETN VPMAMIPNDPLILBRHSOLYEALBELGVLPTIINITIKKRPMAVAGELIYGIV FRFIVYRREBVENFPGIITRGCALCPMIIIGAFINAMAYSRSAYIK"

gene 4101..4952
/gene="HI0905"
/note="HI0905"
/note="similar to GB:J01710 SP:P00470 PID:147987 PID:42685

This Page Blank (usp:0)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:32:25 ; Search time 761 Seconds
(without alignments)
16239.170 Million cell updates/sec

Title: US-09-700-712a-1
Perfect score: 2909
Sequence: 1 gaggaagcttgatgcctc.....cagcgaatcggcgctgcag 2909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_290and04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2909	100.0	2909	3	AAZ40646 Nucleotide
2	1222	42.0	1222	3	AAZ40648 Nucleotide
3	1112.6	38.2	1149	7	ACA53067 Prokaryot
4	842.4	29.0	852	2	AAZ24801 Vibrio ch
5	842.4	29.0	852	7	ACA53066 Prokaryot
6	838	28.8	838	3	AAZ40647 Nucleotide
7	744.2	25.6	110000	2	AAZ40647 Nucleotide
8	505.4	17.4	852	7	ACA42822 Prokaryot
9	462.2	15.9	852	7	ACA44455 Prokaryot
10	403.8	13.9	849	7	ACA44455 Prokaryot
11	292.2	10.0	807	7	AAZ37118 Nucleotide
12	278.2	9.6	110000	7	ACFE7367 27
13	278.2	9.6	110000	7	ACFE5386 0
14	278.2	9.6	110000	7	ACFE5386 1
15	277.8	9.5	876	7	ACFE9784 Phototrab
16	248.2	8.5	876	5	AAH81383 Escherich
17	208.6	7.2	834	6	ADA32523 DNA encod
18	208.4	7.2	834	6	ABG90157 M. capsul
19	161.2	5.5	852	3	AAZ53292 Neisseria
20	156.4	5.4	852	3	AAZ53291 Neisseria
21	156.4	5.4	102634	3	AAH81464 Neisseria
22	156.4	5.4	110000	3	AAH81490 10
23	156.4	5.4	349980	3	AAZ21609 Neisseria

24	152.2	5.2	110000	6	ABA92787_4	Continuation (5 of
25	149.2	5.1	849	7	ABZ40310	Abz40310 N. gonorr
26	149.2	5.1	852	3	AAZ53290	AAZ53290 Neisseria
27	148.6	5.1	1515	4	AAZ61079	AAZ61079 P. putida
28	145.4	5.0	420	6	ABN26154	ABN26154 Human ORF
29	145.4	4.9	96109	4	AAZ28548	AAZ28548 Genomic E
30	142.2	4.9	420	6	ABN26172	ABN26172 Human ORF
31	137	4.7	420	6	ABN26172	ABN26172 Human ORF
32	84	2.9	2440	9	ADBS8174	ADBS8174 Toxicity-
33	84	2.9	2440	9	ADBS8174	ADBS8174 Toxicity-
34	83.6	2.9	2440	9	ADBS8174	ADBS8174 Toxicity-
35	79	2.7	2075	5	AAZ58385	AAZ58385 DNA encod
36	79	2.7	2075	5	ADBS8991	ADBS8991 Toxicity-
37	77.8	2.7	2075	9	ADBS3745	ADBS3745 Primary r
38	75.4	2.6	2308	4	ABA09555	ABA09555 Toxicity
39	75.4	2.6	2208	4	AAK52748	AAK52748 Human Na-
40	75.4	2.6	2280	4	ADBS0585	ADBS0585 Ovarian C
41	75.4	2.6	2288	2	AAV59498	AAV59498 Human bod
42	75.4	2.6	2380	4	AAK51764	AAK51764 Human pol
43	75.4	2.6	2520	4	AAK94823	AAK94823 Human ful
44	75.4	2.6	3952	8	ACH03831	ACH03831 Human CDV
45	75.4	2.6	4137	3	AAZ75484	AAZ75484 DNA encod

ALIGNMENTS

RESULT 1	AAZ40646	standard; DNA; 2909 BP.
ID	AAZ40646	
XX	AAZ40646;	
AC		
XX		
DT	08-MAR-2000	(first entry)
XX		
DE		Nucleotide sequence of V. cholerae thya gene.
XX		
KM		Vibrio cholerae; thya-negative strain; Deltathya; thya gene; vaccine;
XX		cholera; antimicrobial therapy; ss.
OS		Vibrio cholerae.
XX		
PN	MO9961634-A1.	
XX		
PD	02-DEC-1999.	
XX		
PF	21-MAY-1999;	99WO-EP003509.
XX		
PR	26-MAY-1998;	98SB-00001852.
XX		
PA	(SBLV-) SBL VACCIN AB.	
XX		
PI	Carlin N, Lebens MR;	
XX		
PS	WPI; 2000-062719/05.	
XX		
DR	P-PSDB; AAY59126.	
XX		
PT	New Vibrio cholerae strain defective in the thya gene, for use in	
XX		
XX	vacines and for recombinant protein production.	
XX	Claim 9; Fig 1; 42pp; English.	
CC	The invention provides a method for producing a Vibrio cholerae thya-	
CC	negative strain which is Deltathya strain lacking thya gene functions.	
CC	The method comprises site-directed mutagenesis of the V. cholerae	
CC	chromosome to delete and/or insert nucleotides at the thya locus; The V.	
CC	cholerae thya-negative strains are used: for overproduction of	
CC	recombinant proteins; and in vaccines to prevent or treat cholera (or	
CC	other diseases if engineered to express the appropriate proteins). The	
CC	thya gene is also useful for insertion of foreign genes, in a selective	
CC	and site-specific manner, and the proteins expressed by the thya gene or	
CC	by its 5'-flanking region, are useful in research and as targets for	
CC	antimicrobial therapy. When used for recombinant protein production, V.	

CC cholerae provides high yields with secretion of products into the culture
CC medium for ease of subsequent recovery. The thA-negative strain can be
CC maintained by thymine complementation, eliminating the need for
CC antibiotic selection. The present sequence represents the nucleotide
CC sequence of the V. cholerae thA gene
XX

Sequence 2909 BP; 734 A; 644 C; 737 G; 794 T; 0 U; 0 Other;

Query Match 100.0%; Score 2909; DB 3; Length 2909;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 2909; Conservative 0;

```
QY 1 GAGAAAGTTTGTATGCTCAGAGGTTATCTGACAGTTTCCCAATATTTGACCCCGATTTGTT 60
DB 1 GAGAAAGTTTGTATGCTCAGAGGTTATCTGACAGTTTCCCAATATTTGACCCCGATTTGTT 60
QY 61 TTGCAATCGGCCCTTACGCGGTGCGCTGGTATGGCTTATGTTTGGTGGTTTCTTTT 120
DB 61 TTGCAATCGGCCCTTACGCGGTGCGCTGGTATGGTATGTTTGGTGGTTTCTTTT 120
QY 121 TGCTATGTGTGGCCAAATCGCGAGCGGATGCGCGGGCAGTGGTGGAGCGGATGCA 180
DB 121 TGCTATGTGTGGCCAAATCGCGAGCGGATGCGCGGGCAGTGGTGGAGCGGATGCA 180
QY 181 AGTCTCTGACCTTGTATTGCGCGGCTTTTATAGTGTAGTATCGGTGGCCGAGTTGGTTA 240
DB 181 AGTCTCTGACCTTGTATTGCGCGGCTTTTATAGTGTAGTATCGGTGGCCGAGTTGGTTA 240
QY 241 TGTGATCTTTTCAAAATTTTGTATCTGTTCTTGGTCTGACCCCTTTTATTTATCAAAAGTGTG 300
DB 241 TGTGATCTTTTCAAAATTTTGTATCTGTTCTTGGTCTGACCCCTTTTATTTATCAAAAGTGTG 300
QY 301 GACGTGGCGGCAATGCTCTTCCAGCGGGGCTTATTTGGGTGTGATCACCGGCATGTTCTGGTA 360
DB 301 GACGTGGCGGCAATGCTCTTCCAGCGGGGCTTATTTGGGTGTGATCACCGGCATGTTCTGGTA 360
QY 361 TGGCGGTAAAAACAACGCAACGCACTTCTTGGTGTGGCCGATTTTGTGTGCCCTTTTATGTTGC 420
DB 361 TGGCGGTAAAAACAACGCAACGCACTTCTTGGTGTGGCCGATTTTGTGTGCCCTTTTATGTTGC 420
QY 421 ATTGGGTTTGGGAGTGGAGCGTATGCGTATGCTTTATGAAATAGTGAACCTTTGGGGACGAGT 480
DB 421 ATTGGGTTTGGGAGTGGAGCGTATGCGTATGCTTTATGAAATAGTGAACCTTTGGGGACGAGT 480
QY 481 AAOGATGTGCTTGGGCTTTTGTATTCCTTAATGTGTGGCCGACGCGCGCATCTTC 540
DB 481 AAOGATGTGCTTGGGCTTTTGTATTCCTTAATGTGTGGCCGACGCGCGCATCTTC 540
QY 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGTGGCCGACGCGCGCATCTTC 540
DB 541 AACGATGTGCTTGGGCTTTTGTATTCCTTAATGTGTGGCCGACGCGCGCATCTTC 540
QY 541 ACAGCTTTATGAAATTCGCTTATGAAAGCGTGGTCTGTTCTTATTTCTTAAATTTGTTTAT 600
DB 541 ACAGCTTTATGAAATTCGCTTATGAAAGCGTGGTCTGTTCTTATTTCTTAAATTTGTTTAT 600
QY 601 TGTGTAACCTGCTCGCTAGGCGATGCGGATCCGGAAGCTTTTATAGCTGATACGATACATT 660
DB 601 TGTGTAACCTGCTCGCTAGGCGATGCGGATCCGGAAGCTTTTATAGCTGATACGATACATT 660
QY 661 CGGCTTCTGTGTGAAATATGCGGTGAGCGATGCTAGTGGGTCTGTTTGGTGGCT 720
DB 661 CGGCTTCTGTGTGAAATATGCGGTGAGCGATGCTAGTGGGTCTGTTTGGTGGCT 720
QY 721 CATTTCAATGAGGCAAAATCTCTCTTACCTATGATGATCATCGATATTTGATATGAT 780
DB 721 CATTTCAATGAGGCAAAATCTCTCTTACCTATGATGATCATCGATATTTGATATGAT 780
QY 781 TTGGTCTTACAAAGCGGCTTGTATACAGACCGTGTAGACGCAAAATAGGTTAGT 840
DB 781 TTGGTCTTACAAAGCGGCTTGTATACAAAGCGTGTAGACGCAAAATAGGTTAGT 840
QY 841 GAAACAGATTTTATGATCTTTGTACAGGCACTGTCATACAGGTTTGGGTTGAAATATGA 900
DB 841 GAAACAGATTTTATGATCTTTGTACAGGCACTGTCATACAGGTTTGGGTTGAAATATGA 900
QY 901 ACGAACGGGCAAGCGTTGTTGATCTGTGATTAATGCCGATTTGACCTACGATGTGGGCA 960
```

```
DB 901 ACGAACGGGCAAGCGTTGTTGATCTGTGATTAATGCCGATTTGACCTACGATGTGGGCA 960
QY 961 CAATCAGTTTCTCTAGAGCTACACGCAAGTTTTTGAAGCTGCGGACGAGTT 1020
DB 961 CAATCAGTTTCTCTAGAGCTACACGCAAGTTTTTGAAGCTGCGGACGAGTT 1020
QY 1021 GCTCGCTATATTCGTGTTACGATTAATGCGCGGATTTTCCCAATTTAGTATACCAAAAC 1080
DB 1021 GCTCGCTATATTCGTGTTACGATTAATGCGCGGATTTTCCCAATTTAGTATACCAAAAC 1080
QY 1081 CTGGAGTCTATATGCAATTTTAAACCAAGCATGCTCAACATCTTATCCGTAAGGTTGA 1140
DB 1081 CTGGAGTCTATATGCAATTTTAAACCAAGCATGCTCAACATCTTATCCGTAAGGTTGA 1140
QY 1141 GGAATGACATGGGACGCGGTATGAGTTCAGGAGTATGAGCTTGAAGCTGATGTGG 1200
DB 1141 GGAATGACATGGGACGCGGTATGAGTTCAGGAGTATGAGCTTGAAGCTGATGTGG 1200
QY 1201 TCATATTTGACCAAGTTGAAAAAGATTGTGATGATTTGAGCCGTGGCTTGAATGACCGAGG 1260
DB 1201 TCATATTTGACCAAGTTGAAAAAGATTGTGATGATTTGAGCCGTGGCTTGAATGACCGAGG 1260
QY 1261 TGAATTTCTTAACTTTCAAAATTCGCGGTGAATTTCAATGCGGCTTGCCTTGCAT 1320
DB 1261 TGAATTTCTTAACTTTCAAAATTCGCGGTGAATTTCAATGCGGCTTGCCTTGCAT 1320
QY 1321 GTAACGCAATATTTTCAATGCTGGGGGATACCTGTATCTCAACAGTACTACAGCGTTTC 1380
DB 1321 GTAACGCAATATTTTCAATGCTGGGGGATACCTGTATCTCAACAGTACTACAGCGTTTC 1380
QY 1381 ATGTGATGTGCTTGTGGGTTGAATTTCAACATGTGTGAGGTTATGTGTTCTTGGCGCT 1440
DB 1381 ATGTGATGTGCTTGTGGGTTGAATTTCAACATGTGTGAGGTTATGTGTTCTTGGCGCT 1440
QY 1441 GATGGCAAGATCACAGGGAAAAAGCCGGGCTTGGCGTATCAACAGATCTGATGCGCA 1500
DB 1441 GATGGCAAGATCACAGGGAAAAAGCCGGGCTTGGCGTATCAACAGATCTGATGCGCA 1500
QY 1501 CATTTACCAAGATCAACTCGAATTTGATGCGGATGTGAGCTTAAACGTTGACCATTTCC 1560
DB 1501 CATTTACCAAGATCAACTCGAATTTGATGCGGATGTGAGCTTAAACGTTGACCATTTCC 1560
QY 1561 AGCGCTCAGTTCATATCAATCCAAAGATTAAAAACATGAGAGATTGGAACCTTGGGT 1620
DB 1561 AGCGCTCAGTTCATATCAATCCAAAGATTAAAAACATGAGAGATTGGAACCTTGGGT 1620
QY 1621 CACTTGGATGATTTTGAAGTCAACCGGATATCAGTTCCACGATCTTATTTCAATCCCGTT 1680
DB 1621 CACTTGGATGATTTTGAAGTCAACCGGATATCAGTTCCACGATCTTATTTCAATCCCGTT 1680
QY 1681 TTGAGTCAATCCGATATTCAGGGGCTTGTGAGGTTTATATTAATAAAAAAGCTCC 1740
DB 1681 TTGAGTCAATCCGATATTCAGGGGCTTGTGAGGTTTATATTAATAAAAAAGCTCC 1740
QY 1741 CGAAGTGGGAGCTTTTATACAGATGATGCTTTAAGCCTTAAAGGTTAGGGCAAG 1800
DB 1741 CGAAGTGGGAGCTTTTATACAGATGATGCTTTAAGCCTTAAAGGTTAGGGCAAG 1800
QY 1801 AATGCTGCGGGGATGACGACAAACACCCCAATATGATACACACCACTTTTCT 1860
DB 1801 AATGCTGCGGGGATGACGACAAACACCCCAATATGATACACACCACTTTTCT 1860
QY 1861 CTTAACAACCCCAAGTTGAGTGAAGTGTGACCTTTAATAGGCGTTGCGGTTAAGAAAG 1920
DB 1861 CTTAACAACCCCAAGTTGAGTGAAGTGTGACCTTTAATAGGCGTTGCGGTTAAGAAAG 1920
QY 1921 AATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAAGTGCACACAGCGCAATTGG 1980
DB 1921 AATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAAGTGCACACAGCGCAATTGG 1980
QY 1981 CAGAGCAAAACAGGCAAACTCAACAGACAGCGGTTCCGGCAGACAGACGATATATCA 2040
```



```

Db      421 GGTGAATCTTAACCTTACAAATCGGGTGAATTCACATGGGGTGTTCGCCCTTGC 480
QY      1319 ATGTACAGCAGCATCTTTTCATTCGTGGGGGATACCTTGATCTCAACAGTACTACGGCT 1378
Db      481 ATGTACAGCAGCATCTTTTCATTCGTGGGGGATACCTTGATCTCAACAGTACTACGGCT 540
QY      1379 TCATGTGATGTGCTTGGGGGTGAATTTCAACATGTGAGGTTATGTGTTCTTGGC 1438
Db      541 TCATGTGATGTGCTTGGGGGTGAATTTCAACATGTGAGGTTATGTGTTCTTGGC 600
QY      1439 CTGATGCAACATCAACAGGAAAAAGCCGGCTTGGCTATCAAAATCGTCAATGCG 1498
Db      601 CTGATGCAACATCAACAGGAAAAAGCCGGCTTGGCTATCAAAATCGTCAATGCG 660
QY      1499 CACATTACCAAGATCACTCCAAATGTAGTGGCGATGTGACGTAAAGTGGCCATTC 1558
Db      661 CACATTACCAAGATCACTCCAAATGTAGTGGCGATGTGACGTAAAGTGGCCATTC 720
QY      1559 CCAGCGCCTCAGTTCATATCAATCCAAAGATTAAACACTGACAGATTGGAACTTGG 1618
Db      721 CCAGCGCCTCAGTTCATATCAATCCAAAGATTAAACACTGACAGATTGGAACTTGG 780
QY      1619 GTCACTTTGATGATTTTGAAGCTCACCGGATATCACTTCACGATCTTAATACCG 1678
Db      781 GTCACTTTGATGATTTTGAAGCTCACCGGATATCACTTCACGATCTTAATACCG 840
QY      1679 TTTTCAGTCTAA 1690
Db      841 TTTTCAGTCTAA 852

```

RESULT 6

AA240647
ID AA240647 standard; DNA; 838 BP.

AA240647;

08-MAR-2000 (first entry)

Nucleotide sequence of V. cholerae *thyA* gene 5' flanking region.

Vibrio cholerae; *thyA*-negative strain; *Delta*thyA; *thyA* gene; vaccine;
cholera; antimicrobial therapy; ss.

Vibrio cholerae.

WO9961634-A1.

02-DEC-1999.

21-MAY-1999; 99WO-EP003509.

26-MAY-1998; 98SE-00001852.

(SBLV-) SBL VACCIN AB.

Carlin N, Lebens MR;

WPI; 2000-062719/05.

P-PSDB; AAY59127.

New Vibrio cholerae strain defective in the *thyA* gene, for use in
vaccine and for recombinant protein production.

Claim 10; Fig 2; 42p; English.

The invention provides a method for producing a Vibrio cholerae *thyA*-
negative strain which is *Delta*thyA strain lacking *thyA* gene functions.
The method comprises site-directed mutagenesis of the V. cholerae
chromosome to delete and/or insert nucleotides at the *thyA* locus. The V.
cholerae *thyA*-negative strains are used: for overproduction of
recombinant proteins; and in vaccines to prevent or treat cholera (or

other diseases if engineered to express the appropriate proteins). The
CC *thyA* gene is also useful for insertion of foreign genes, in a selective
CC and site-specific manner, and the proteins expressed by the *thyA* gene or
CC by its 5'-flanking region, are useful in research and as targets for
CC antimicrobial therapy. When used for recombinant protein production, V.
CC cholerae provides high yields with secretion of products into the culture
CC medium for ease of subsequent recovery. The *thyA*-negative strain can be
CC maintained by thymine complementation, eliminating the need for
CC antibiotic selection. The present sequence represents the nucleotide
CC sequence of the V. cholerae *thyA* gene 5' flanking region

SQ Sequence 838 BP; 145 A; 176 C; 232 G; 285 T; 0 U; 0 Other;

Query Match 28.8%; Score 838; DB 3; Length 838;

Best Local Similarity 100.0%; Pred. No. 4.4e-256;

Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAAAGTTTGTATGCTCAGGGTTATCTGACATTTCCCAATATTGACCCCGTATTGTT 60

Db 1 GAGAAAGTTTGTATGCTCAGGGTTATCTGACATTTCCCAATATTGACCCCGTATTGTT 60

QY 61 TTGATCGGCTCTAGCGGTGCGCTGTATGCTTGAATATTGGTGGTTTCTTTT 120

Db 61 TTGATCGGCTCTAGCGGTGCGCTGTATGCTTGAATATTGGTGGTTTCTTTT 120

QY 121 TGTATGTGTTGGGCAATCGCGGACGGATGCGCGGCAATGTTGGACGCGTGAACA 180

Db 121 TGTATGTGTTGGGCAATCGCGGACGGATGCGCGGCAATGTTGGACGCGTGAACA 180

QY 181 AGTCTCTGACTTGTATTTGCGCGGCTTTTATGATGATGATGATGATGATGATGATGAT 240

Db 181 AGTCTCTGACTTGTATTTGCGCGGCTTTTATGATGATGATGATGATGATGATGATGAT 240

QY 241 TGTGATCTTCTACATTTTGAATGCTGCTTGTGCTGACCTCTTTATTTTAAAGTGTG 300

Db 241 TGTGATCTTCTACATTTTGAATGCTGCTTGTGCTGACCTCTTTATTTTAAAGTGTG 300

QY 301 GACGCGGCGCATGCTCTTCCACGCGGCTTATGATGATGATGATGATGATGATGATGAT 360

Db 301 GACGCGGCGCATGCTCTTCCACGCGGCTTATGATGATGATGATGATGATGATGATGAT 360

QY 361 TGCAGGTAACCAACGACCTTCTTGTGTTGCGGCAATTTGTTGCGCTTTAGTGGC 420

Db 361 TGCAGGTAACCAACGACCTTCTTGTGTTGCGGCAATTTGTTGCGCTTTAGTGGC 420

QY 421 ATTGGTTTGGGATGGAAGTATCGTACTTTATGAATGATGAATTTTGGGACGAGT 480

Db 421 ATTGGTTTGGGATGGAAGTATCGTACTTTATGAATGATGAATTTTGGGACGAGT 480

QY 481 AACGATGCGCTTGGGCTTTTGTATTCCTTAATGATGATGATGATGATGATGATGATGAT 540

Db 481 AACGATGCGCTTGGGCTTTTGTATTCCTTAATGATGATGATGATGATGATGATGATGAT 540

QY 541 ACAGCTTATGATATGCTTGAAGGCGTGTCTGTTCTTATTTCTTAATTTGTTTAT 600

Db 541 ACAGCTTATGATATGCTTGAAGGCGTGTCTGTTCTTATTTCTTAATTTGTTTAT 600

QY 601 TGTAAACCTGTCGCTAGGACGCTATCGGACTGTTTATGCTGATACGATACAT 660

Db 601 TGTAAACCTGTCGCTAGGACGCTATCGGACTGTTTATGCTGATACGATACAT 660

QY 661 CCGCTTCTTGTGGAATACGTCGAGCAAGATCTCAAGTGGTGTGTTGGTGGCT 720

Db 661 CCGCTTCTTGTGGAATACGTCGAGCAAGATCTCAAGTGGTGTGTTGGTGGCT 720

QY 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATGAT 780

Db 721 CATTTCAATGGGCAATCTCTCTTACTATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 TTGGTCTTCAACGCGGCTTGTATCAAGACGCTGTAGACAGCAATATGGGTAGTTAG 838

Db 781 TTGGTCTTCAACGCGGCTTGTATCAAGACGCTGTAGACAGCAATATGGGTAGTTAG 838

```
RESULT 7
AAT42063_09
Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete ge
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1830121

Query Match 25.6%; Score 744.2; DB 2; Length 110000;
Best Local Similarity 66.7%; Pred. No. 7.9e-225;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

OY 2 AGAAGTTTGTATGCTCAGGGTTATCTGCAGTTTCCCAATATTTGACCCGATATGTTT 61
DB 58086 ATAAGGAAATATATGAATTCAAATATTTACTTCTTCCCACTTGATCCGAGATTTT 58145
OY 62 TCGATCGC-----CTCTAGCGGTGGCGGTATGCTGATATTTGGTGGTTTC 115
DB 58146 ACGCTTGCAGATGTAATATCGGCTTACGTTGTAATGCTGATGATACCTTTAGGTTT 58205
OY 116 CTTTTCGTATGTGTGGCCATTCGCCAGCGAGTCGCGCGGAGAGTGTGACGCGT 175
DB 58206 GTTTTTCACGTGCTGCGGTTGCGGTTGCGGTGTAATCCGCCAAATAGCGGTTGACAGTA 58265
OY 176 GAGCAAGTCTGACTGTGATTTGCGCGGCTTTTAAAGTATGATCGTGGCCAGTT 235
DB 58266 GATCAAGTTGATAGCTTACTTTTCAACGTTTATGGGGGTGTTATTTGGCGAGTGT 58325
OY 236 GGTATGATGATCTTCAATTTTATCTGTTCTTCTGACACCCCTTATTTATTCAAA 295
DB 58326 GCGAGTGTATTTCTATATCTCGATCATTTTCTTACAAGAACCACTTTATTTATTCGCG 58385
OY 296 GTGTGACTGGCGCATGCTCTTCCACGCGGCTTATTTGGGTGATCACGCGCATGTT 355
DB 58386 GTTTGCGAAGTGTGATGTCGTTCCACGCGTCTTAATTTGGTATTTGGTGTATTT 58445
OY 356 TGTATGCGCGTAAAAACCAACGACCTTCTTGTGTGGCCGATTTTGTGCCCTTTA 415
DB 58446 TGGACATCTTATTTCTCAAAAAGTATTTTGGCAACGCGCTGATTTTGTGCGCTTTG 58505
OY 416 GTGCGATCTGTTTGGGAGTGGAGTATCCGTACTTATGAAATAGGAATCTTTGGGA 475
DB 58506 ATTCGTTTGTGTTAGGTTTAGGAGAAATGTGTAATTTCAATTAATCTTGAACATAAGGGA 58565
OY 476 CGAGTACGAGATGTCCTTGGGCTTTTGTATTCCTTAATGTTGGCC--CACTGCGCGGC 532
DB 58566 CCGCAAGGATGTGCTTGGCAATGATTTTCCGAATGATCTCTTTTACTGCTCTGT 58625
OY 533 CATCTTCACAGCTTATGAAATTCGCTTAGAAGCGTGTCTTCTTATTTCTTAT 592
DB 58626 CATCATCAACAATTATGAAGCTTTTAAAGGCGTGTGTGTTTACGATTCGAT 58685
OY 593 TGTATTATGTAACCTGCTCGCTAGGACGCTATCCGAGCTTTTATAGCTGGAATC 652
DB 58686 ATTTTATTAATAAACCAAGTCATGCTTCTGTTGAGGTTTATCTTATTTGTTAT 58745
```

```
OY 653 GGTACATTTCCGCTTCTTGTGAATACGTCGAGACCGAGATGCTCACTTGGCTGTGTT 712
DB 58746 GCGGCTTCCGTTTATTTGTGAATATGTGTGTAACCTGAAG-----TTGAAATTTTC 58799
OY 713 GGTGCTTCAATTTCAATGGGCAATCTCTCTTACTTATGATGATCATCGTATTTTG 772
DB 58800 TTTGGGATTTATTAACGAGGGCAAGCCCTTGTGCGATGATTAATTTGTGTGCTTTC 58859
OY 773 ATGATGTTTGTCTTACAAAGCGGTTTGTATCAAGACCGTATGACGAAATAGGGT 832
DB 58860 ATTTATGCTTGGCTTATTCACG-----AAAGTCCGCTATTAATATAGAG 58907
OY 833 AGTTAGTGAACAGATTTTATGATCTTTGTGACGCGCATCTGTCAATCAAGTGTGGTT 892
DB 58908 ATTTTATGAAGCAATATCTTGAAGTGTGCTGCGCATTTTATGTAAGGGAAATGGGTT 58967
OY 893 GAAATGAACGAACGGGCAACCGTTGTTGACTGTGATTAATGCGGATTTGACCTACAT 952
DB 58968 GCTATGAACGTACAGGTACATTTGCTCACGGTCATTAATGAGATTTTGAATATGAT 59027
OY 953 GTGGCAACAATCAAGTTTCTCTAGTGACTACAGCAAGATTTTGGAAAGCTGCCGTA 1012
DB 59028 GTGGCAATATCAATTTCCGCTGATTACTACCCGTAAAGTTATTTGAAAGCGCGCAT 59087
OY 1013 GCGAGTTGCTCGGCTATATTCGTGTTACGATATGCGGAGATTTTCGCAATTAAGT 1072
DB 59088 GCTGAATTTTATAGTTATATTCGTGATATGACATATGCGCTGATTTCCGCGCATGGC 59147
OY 1073 ACCAAACCTGGATGCTAATGCAATTTTAAACCAAGATGCTTCACATATCTTACGT 1132
DB 59148 ACGAAATCTGGAGTGTCTAATGCGAATGAATAATGCAAGTTGCGCTTTCGAAATCCGATCGT 59207
OY 1133 AAAGTGAGATGACATGGAACGCGTATGATGTTGCGAGTGAAGCTTGAAGCTTAAGCT 1192
DB 59208 AGAGCGTTGATGATATGAGGCGCGGTATATGTGTGCAAGCAAGCATGCGTATAGCT 59267
OY 1193 GATGTGTCTATATTTGACCAAGTGAATAAGATTTGTATGATTTGAGCGCGTGTGAT 1252
DB 59268 AATGGAATACTATCGATCAGCTACGTAAATTTGTATATCTTAAACGAAGTATGAT 59327
OY 1253 GACCGAGTGAATTTCTTAACTTCTACAAATCCGGGTGAATTTACATGCGGTGTTGGC 1312
DB 59328 GATAGAGAGATTTTAACTTTTAACTTCCGCGGAATTTGATCTTGTGTGCTTCGT 59387
OY 1313 CTTGCAATGACAGCAATCTTTTCAATGCTGCGGAGATACCTGTATCTCAAGTACT 1372
DB 59388 CTTGTATGATGATGCAATACATCTTTTCTTGTGGGCAATCTTACATCTTACATGCTAT 59447
OY 1373 CAGGCTTCAATGATGTCCTTGGGTTGAATTTCAACATGTCAGAGTTTATGTGTC 1432
DB 59448 CAGGCTTCTGTGATGTTCCGCTTGAATTTCAATCAATTTCAAGGTGTTTACCTTC 59507
OY 1433 CTTCGCGTATGAGCAAGATCAACAGGAAAGACCGGCTTGGGATCAACAATGCTC 1492
DB 59508 TTACACTTATGAGCTCAAGTACGCGGCAAAAACCGGCAAGGATATCATTAATATGTG 59567
OY 1493 AATGCGACATTTTCAAGATCAACTGAAATTTGATGCGCATGTCAGCTAAAGCTGAG 1552
DB 59568 AATGCGCATTTTATGAGATCAAGCTTGAATTTATGCTGACGTACACATTTAAACGCGAG 59627
OY 1553 CCATTTCCAGCGCTCAAGTTCATATCAATCAAGATTTAAACATGACAGATTTGGAA 1612
DB 59628 CTTTCCCATTAACCAAAATTTGAATTTATCCAGATTAATAAAGCTTGAAGATTTGGAA 59687
OY 1613 ACTTGGCTCACTTGAATGATTTTGAAGCTCACCGGATATCACTTCCACAGATCTTAA 1672
DB 59688 ACTTGGGTCAGAGATGATTTTAAAGTCTTGGCTATCAATCCACGAACCAATTA 59747
OY 1673 TACCGTTTCAAGCTAAT 1691
DB 59748 TATCCTTTTCGCTAAT 59766
```

RESULT 8
ACAA2822
ID ACAA2822 standard; DNA; 852 BP.
XX
AC ACAA2822;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #24479.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Pasteurella multocida.
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362659P.
XX
PA (ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR MPI; 2003-029926/02.
XX
P-PSDB; ABUJ8952.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 30692; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence date for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 852 BP; 247 A; 201 C; 181 G; 223 T; 0 U; 0 Other;
Query Match 17.4%; Score 505.4; DB 7; Length 852;
Best Local Similarity 74.6%; Pred. No. 6.8e-150;
Matches 635; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
QY 840 TGAACAGATTTAGATCTTTGTGTCAGCGCATGCGATCAAGATGTTGGTTGAAATG 899
DB 2 TGAACACATTTTAGACTTTGCCAAGCATGATGTTGATGAAGTGTGGTTGAAATG 61
QY 900 AACGAACGGGCAAGCGTTGTTGACTGTGATTAATCCGATTTGACCTACGATGCGCA 959
DB 62 CAGGACTGCGCAACGTTGCTGACCGATTTGATGCGGATCTCACTACGATGTTGCCA 121
QY 960 ACAATCAGTTTCTCTAGTACATCAACGCAAGATTTTGGAAAGTGCCTGACCGACT 1019
DB 122 ATTAACCAATTTCCGTTGATCACACACGCAAAAGCTATTGAAAGCGCAATTGCAAT 181
QY 1020 TGCTCGGCTATATTCGTTGATGATGATGCGGCGGATTTTGCCCAATTAGGTACCAAA 1079
DB 182 TTTTAGGTTATATCCGTGCTATGACAAATGCGGCAATTTCCGCAACTCGGACCAAAA 241
QY 1080 CCTGGAGTCTATATGCCAATTTAAACCAAGCATGCTCAACATCTTAACGTTAAAGTG 1139
DB 242 CCGGAGTCCCATGCTATGAAACAGTGCATGCTCAATATCCGACCGCAAGGCA 301
QY 1140 AGATGACATGCGGACCGCTGATGCTGTTCAAGGTAGAGCTTGGGCTAGCCTGATGTC 1199
DB 302 CCGATGATATGGAAGCGTGTATGCGGTAACAGGCGGTGCAATGCGCAACCGAAGCGTG 361
QY 1200 GTCATATGACAGTTGAAAGATTTGATGATTTGAGCCGTGCGTTGATGACCGAG 1259
DB 362 AAACGTCATATTAATTCGCAAAATGCTCAATATTTGCGCAACGGCATTTGACCGTG 421
QY 1260 GTGAATTTCTTAATCTTCAAAATCCGGTGAATTTTCAATGCGGATTTTGCCTTGA 1319
DB 422 GTGAATATATGACCTTTCAACCCCGGATTTGAATGAGGTTCCATGACCTTGA 481
QY 1320 TGTACAGCATATTTTTCATTTGCTGCGGAGTACCTTGTATCTCAACAGTACTGACGCTT 1379
DB 482 TGCACACCCACACCTTCTCATTAATCTCGGTGATACCTCTATCTCAACAGTTATCAGCGTT 541
QY 1380 CATTGATGTCCTTGGGTTGATTTCAATGATGATGAGGTTATGTTCTTGGCG 1439
DB 542 CTGTGATGTCCTGCTGCTGTTGAATTTCAACAAATTCAGTGTACCTTTCTTGCT 601
QY 1440 TGATGCAACAGATCAAGGCAAAAGCCGCGCTTGGCTGATCAAGATGCTCAATGCGC 1499
DB 602 TGATGCAAAATTTACGGGCAAAAAGCGGTAAAGCGTAAACAAATCATCAATGAC 661
QY 1500 AATTTTACCAAGATCACTGATGATGCGGATGCGGATGCAAGCTTAAAGCTGACCTTC 1559
DB 662 AATTTTATTAAGATCAACTGATCTCATGAAATTAAGTGAATTTAAACCTGACCTTCC 721
QY 1560 CAGCGCTTCAGTTCCATATCATCCAAAGATTTAAACCTGAGATTTGAAACTTGGG 1619
DB 722 CGTTGCACAATTTACCATTAACCCCGAATTTAAACTCTGAAAGATTTGAAACTTGGG 781
QY 1620 TCATTTGGATGATTTTGAAGTCAACCGGATTCAGTTCCAGATCTTATTCATACCCGT 1679
DB 782 TGACATATGATATTTTAAGTACAGGCTATCAATGCAACATGCAATTAATATCTT 841
QY 1680 TTTCAAGTTAA 1690
DB 842 TCTCTGTGTA 852
RESULT 9
ACAA4292
ID ACAA4292 standard; DNA; 852 BP.
XX
AC ACAA4292;

XX	drug design; gene.
XX	
OS	Proteus sp.
XX	
PN	WO200277183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002MO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
XX	
PR	06-SEP-2001; 2001US-00948993.
XX	
PR	25-OCT-2001; 2001US-0342923P.
XX	
PR	08-FEB-2002; 2002US-00072851.
XX	
PR	06-MAR-2002; 2002US-0362699P.
XX	
PA	(ELIT-) ELITRA PHARM INC.
XX	
P1	Mang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KU, Zykkind JW;
XX	
P1	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
XX	
DR	P-PSDB; ABU40585.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
XX	
PT	for homologous nucleic acids required for cellular proliferation to
XX	
PT	isolate candidate molecules for rational drug discovery programs.
XX	
PS	Claim 14; SEQ ID NO 32325; 1766bp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising any one of
XX	
CC	the 6213 antisense sequences given in the specification where expression
XX	
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
XX	
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
XX	
CC	encoding a polypeptide whose expression is inhibited by the antisense
XX	
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX	
CC	polypeptide or its fragment whose expression is inhibited by the
XX	
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
XX	
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX	
CC	proliferation or the activity of a gene in an operon required for
XX	
CC	proliferation; (7) identifying a compound that influences the activity of
XX	
CC	the gene product or that has an activity against a biological pathway
XX	
CC	required for proliferation, or that inhibits cellular proliferation; (8)
XX	
CC	identifying a gene required for cellular proliferation or the biological
XX	
CC	pathway in which a proliferation-required gene or its gene product lies
XX	
CC	or a gene on which the test compound that inhibits proliferation of an
XX	
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX	
CC	compound's activity; (11) a culture comprising strains in which the gene
XX	
CC	product is overexpressed or underexpressed; (12) determining the extent
XX	
CC	to which each of the strains is present in a culture or collection of
XX	
CC	strains; or (13) identifying the target of a compound that inhibits the
XX	
CC	proliferation of an organism. The antisense nucleic acids are useful for
XX	
CC	identifying proteins or screening for homologous nucleic acids required
XX	
CC	for cellular proliferation to isolate candidate molecules for rational
XX	
CC	drug discovery programs, or for screening homologous nucleic acids
XX	
CC	required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> ,
XX	
CC	<i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence data for this patent did
XX	
CC	not form part of the printed specification, but was obtained in
XX	
CC	electronic format directly from WIPO at
XX	
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SO	Sequence 849 BP; 263 A; 155 C; 189 G; 242 T; 0 U; 0 Other;
XX	
Query Match	13.9%; Score 403.8; DB 7; Length 849;
XX	
Best Local Similarity	67.3%; Pred. No. 1.9e-117;
XX	
Matches 570; Conservative	0; Mismatches 277; Indels 0; Gaps 0
XX	
840	TGAACAGATTTAGATCTTTGACAGCGATCGTCATCAAGGTTTGGGTTGAAGAAG 899
XX	
DB	2 TGAAAGCATGTTGCGCATTTGTCTCAACGCATTTATCATGAAGGTGAATGATTGATTA 61
XX	
900	AACGAACGGGCAAGCGTTTGATCGTGTATTAATCGATTTGACCTACGATGATGGCA 959
XX	

D	b		62	AACGGAAGGAAACCCTGTGGTTTAACGGTGATTAATGCCATCTGGAAATATGATGGCCAA	121
O	y		960	ACAATCAGTTCCTCTTAGTACTTACACGCAGAAGTTTTGGAAAAGCTGCCGTACCGAGT	1019
D	b		122	ATAACCAATTCCACTCATTAACAACGGGTAAAAAGTTTTTCAAAGCCGCATTGCCGAMC	181
O	y		1020	TGCTCGGCTATATTCGTCGGTTAAGATTAATGCGCGCATTTTTCGCCATTATGATACCAAA	1079
D	b		182	TATTAGGGTACTTAACGGTGGTTATGATATATCAGACAAATTCGCTATATGGCTGTATATA	241
O	y		1080	CCTGGAGTGCATATGCAATTTTAAACCAACATATGCTCAACAATCCTTACCGTAAAGTG	1139
D	b		242	GTTGGAACGGGAATGCTAATGAAAAATAGCGCATGGTTAATTAATCCGCAATCGTAAAGGTG	301
O	y		1140	AGATGACATAGGACGCGGTATATGTTTCAGGGTAGAGCTTGGCTTAAGCTTGATGGTG	1199
D	b		302	AAGATGATATGGGCGAGTCTATATGTTTCAAGGACGCCAATGGCAACGCCCTGATGGCT	361
O	y		1200	GTCATATTTACACAGTTAAAAAAGTTTGTATGATTTTGAACCGTGGCGCTGATGACCGAG	1259
D	b		362	CGCCTTTGATTCATATTACGTAAGTGTATATTAATCTTAACATAACGATATTAAGATGGTG	421
O	y		1260	GTCGAATTTCTAACCTTCAATCCGGGTAAATTTCAATAGGGAGTGTTCGGCCCTTGCA	1319
D	b		422	GTCGATAGTCACTTTTATTAATATCCGGAAGAAACAGCCTTAGGGTGTTAACGTCAATGCA	481
O	y		1320	TGTACAGCCATCATTTTTCATTTGCTGGGGATACCTTGATCTCAACAGTACTCACGCTT	1379
D	b		482	TGCATACACATCTTTTCAATGCTCGGTGACAGGCTTATTTAATCCTCATATCAAGTAA	541
O	y		1380	CATGTGATGTCCCTTGGGTTGAAATTTCAATGTGTGACAGTTTATGTGTCTTCTTGGC	1439
D	b		542	GCTGTGATGTCCCATTAGGGTGTGAATTTTATCAAAATTCATGTCTTGTATTTATAGCGC	601
O	y		1440	TGATGGACAGATACAGGAAAAAGCCGGGCTTGGCGTATCACAAGATGCTCAATGCGC	1499
D	b		602	TGGTGCTCAATACACGGGCAATPAACAGAGTAAGCATTCACATPAAGATATGTTAATGCTC	661
O	y		1500	ACATTTACCAAGATCACTCGAATTGATGGCGATGTGACAGTAAACGTAGCCATTCC	1559
D	b		662	ACATTTATGGAACCAATTACCTGTATGTGTATGTTTCAACTAAAAAGAGCCATTGC	721
O	y		1560	CAGGCGCTCAGTTCATATCAATCCAAGATTTAAAAACATGACGAGATTGGAAACTTGGG	1619
D	b		722	CATTACCTAAGCTACATATTATTCAAAACATTAAGACATTAGATGATAGAAACTTGGG	781
O	y		1620	TCACTTTGATGATTTTGAAGTCAACCGGATATCAGTTCACGATCCTATTCATACCCGT	1679
D	b		782	TCAAGACAGATGACTTATGTTGAAAGTATACAGTGTCAAGAGCCGATAAATACCCGT	841
O	y		1680	TTTCAGT 1686	
D	b		842	TCACGGT 848	
RESULT 11					
AAZ37118					
ID	AAZ37118 standard; DNA, 807 BP.				
XX	AAZ37118;				
XX	28-JAN-2000 (first entry)				
DE	Nucleotide sequence of H.influenzae HI0904 gene.				
XX	Genome; mutagenesis; transposon; isolate; locate; essential gene; ss;				
KW	detect; growth; anti-microbial therapy; genomic footprinting.				
XX	Haemophilus influenzae.				
OS					
XX					
FN	W09J50402-A1.				

Db	104168	ATGAGTAAACAGCTACCTGGCAATTCCTAATATATGACCCGGTCATTTTTCATATAGTCCA	104227
Qy	74	CTAGCGGTCCGCTGGTATAGCTTGAATATTTGGTGGGTTTCCCTTTTGTATATGGTTG	133
Db	104228	ATAGCCCTCCCATTTGGTATATGGTTTCAATGTACTTGGTGGGTTTGTATTCGCCAATGGGCTA	104287
Qy	134	GCCAAATCCCGACGAGATCGCCGGCAGTGGTGGACCGGTAGACAAATCTCTGACTTG	193
Db	104288	GCGACACCGAGGGCAGCCGAAACCAATAGTGGTTGGACAAAACGAACTTAGAACTTG	104347
Qy	194	TTATTCGCCGCTTTTATAGTGTAGTATCGGTGGCCGAGTTGGTTATGTATCTTCTAC	253
Db	104348	CTTATGACAGTTTGTCTGGCCATATTTGTGTGGCCGCTCGTGGCTACACCTCTGTTTAT	104407
Qy	254	AATTTTGAATCTGTCCCTGGTGCACCCCTTTTATTTATCAAAAGTGGACGTGGCGGCTAG	313
Db	104408	AAATTTCCCTGCTCTTCTGCACAAACCCCTTTATTTTGTAAAGTTGGAGATGGGAGATG	104467
Qy	314	TCCTTTCACGCGCGCTTATATGGGTGTATCACCGCAGTTCTGTATSCGGGTAAAC	373
Db	104468	TCCTTTCACGCGGTTTATAGTGGGGTTATCTCGCAATGTGGTGTTTGGCCGCAACAG	104527
Qy	374	CAAGCACCTTCTTTGGTGTGGCCGATTTTGTGCCCTTTATGTGCATTGCGTTGGGG	433
Db	104528	AAACCCCATTTCTTTCAGATGCTCACTTCATTTGCCCGCTTGTGTGCTTTTGTATAGC	104587
Qy	434	ATGGACGTAATCGGTAACTTTATGAATAGTGAACTTTGGGACGAGTAAC---GGATGTG	490
Db	104588	ATGGGGAATATCGTATATTTTATTAACGCGAGTTGTGGGACGTCTCACATTGGATAG	104647
Qy	491	CTTGGGCTTTTGTATTCCTTA-----	512
Db	104648	CCGTGGCGATGTATTTCCCAAGCTCCCGTGGTAAGATATTTGATTCATGGCTGCACCGAT	104707
Qy	513	-----ATGGTGCCCATCTGCCCGGCATCTTCACAGCT	547
Db	104708	CCCTCAGTGTATCAGTACTTGAGCAATATGGTGTCTGCCAGTACCTTCACAGCTT	104767
Qy	548	TATGAATTCGCTTGAAGGCGGTGTTCTGTTCTTATTTCTTAATGGTTTATTTGGTAA	607
Db	104768	TATGAATATGGCGTTGGAAGAGATCGTCTGTTCAATATTTTGAATTTATATATCCGTAA	104827
Qy	608	CCTGTCCGCTAGGACGCTATCCGGACGTGTTTATAGCTGATATCGGTACATTCGCGCTTC	667
Db	104828	CCTCGCCCAATGGGACAGTGTTCGGGTTATTTCTTAATTTGGCTACGCTAATTTTCCAGTG	104887
Qy	668	CTTGGGAATAGCGTCGATGACGAGATGCTACATTGGGCTGTGGTGGCTTCATTTCA	727
Db	104888	ATAGTCGAGTTTTCGCCCAACCAATGACCAACTGGCTTATTCGACCGGATCAT	104944
Qy	728	ATGGGCAATCTCTCTTACCTATATGTATGTATCATCGTATTTTGAATGTTGGTCT	787
Db	104945	ATGGGCAAAATCTTTCTATACCAATGATCTGGCGGGAATATTAATGATGATATGGCA	105004
Qy	788	TACAAGCGCGTTTGTATCAAGCGTATGACGAAATATGGGTGTATG	838
Db	105005	TATAAACCAAGTATATAGTATCAAGGATTAATGAACAGTATCTGG	105055

Fragment Name	Begin	End	LOCUS	ACF65386	Accession	ACF65386
ACF65386_0	1	110000				
ACF65386_1	100001	210000				
ACF65386_2	200001	310000				
ACF65386_3	300001	410000				
ACF65386_4	400001	510000				
ACF65386_5	500001	610000				
ACF65386_6	600001	707779				

	Best Local Similarity	61.1%	Pred. No. 4.8e-76f;	
Matches	544;	Conservative	0;	Mismatches 278; Indels 69; Gaps 3;
OY	14	ATGCCCTAAGGTTATCTGCAAGTTTCCAAATATTGACCCCGATATGTTTGATGGGCCCT	73	
Db	4168	ATGAGTAAACAGCTACCTGGCAATTTCCAAATATTGACCCCGCATTTTTCAAATAGTGTCA	4227	
OY	74	CTACCGGAGCGCTGGTATGAGCTTGATATGATTTGATGGGTTTCTTTTGTGATATGGTGG	133	
Db	4228	ATAGCCCTCCATTGGTATGGTTTATGTTATGTTATCTTGGTTGGGTTTGTATTCGCAATGTGCTA	4287	
OY	134	GCCAAATGCGCGAGCGGATTCGGCGGGGAGTGGTTGACGCGGTGAGCAAGTCTTGACTTG	193	
Db	4288	GCGACAGCAGGCGGAGGAAACCCAAATATGTTGGTGGACAAAAAACGAAGTAGAGAACTTG	4347	
OY	194	TTATTCGCGGCTTTTATGAGTGTATGATCGGTGGCCGAGTTGGTTATGTGATCTTCTAC	253	
Db	4348	CTTATATCAGGTTTGTGCTGGCGTATTTGTGGTGCGGTCTGGGCTACGTCTGTTTAT	4407	
OY	254	AATTTATATCTGTCTCTTGCGACCCCTCTTATTTATTCAAAGTGTGACTGGCGGANTG	313	
Db	4408	AATTTCCCTGCTTTCTTGACAAACCCCTTTATTTTATTTAAAGTTTGGATGTGGGATG	4467	
OY	314	TCTTTCACGCGCGGCTTATTTGGTGTGATGACGCGCAATGTTCTGTATGCGGTAAAC	373	
Db	4468	TCTTTCACGCGGCTTATTTAGTGGGGTATATCTGGCAATGTGTGTGTTGGCGGAGAACG	4527	
OY	374	CAACGCACCTTCTTTGGTGTGGCCGATTTTGTGCCCTTTAGTCCATTGGTTTGGGG	433	
Db	4528	AAACGCATTTCTTACAAAGTTGCTGACCTTACCTCCCGTTAAGTCCCTTTGGGTTTAAAGC	4587	
OY	434	ATGGAGCGTATCCGTTACTTTATGAAATATGAACTTTGGGGAGAGTAAAC---GGAATGTG	490	
Db	4588	ATGGGAGAAATCGGTAAATTTTATTAACGGCGAGTTGTGGGACGTGTCACTTGGATACG	4647	
OY	491	CCTTGGGCTTTTGTATTCCTTA-----	512	
Db	4648	CCGTGGCGATGTTATTTCCCAAGTCCCGTGTGAATATATGATTTGGCTGCAACCGAT	4707	
OY	513	-----		
Db	4708	CCCTCAGTGTATCAGTACTTGAGCAATATGTTCTGTGCAACGTCAACCTTCAACACTT	4767	
OY	548	TATGATATTCGCTTGAAGAGCGTGTCTGTCTTTATTTCTTATATGTTTATTTGTAA	607	
Db	4768	TATGAAATATGGGTGGAAGGAGTGTCTGTCTTCAATATTTGAATTTATATATATCCGTAA	4827	
OY	608	CCTGATCGGATGAGCAGGTATCCGGAAGTGTTTTACCTGGATACGGTACATTCGCGTTC	667	
Db	4828	CCTGCGCCCAATGGGCAAGTGTTCGCGTTTATTTCTTAATTTGGCTACGGTATTTTCGAGTG	4887	
OY	668	CTTGTGAATATCGTCCGTGAGCCAGATGCTCAGTTGGGTCTGTTTGGTGGCTTCATTCA	727	
Db	4888	ATATGTCAGTTCCTTCCGCAACAGATGCAAACTGGGCTTATTCGACGGGATCAGT---	4944	
OY	728	ATGGGCGAAATCTCTCTCTTACCTATATGTATATCTGGTATTTTGAATATGATGGTTGGTCT	787	
Db	4945	ATGGGCGCAAAATCTTCTTATATCCATATATCTGGCGGGAATTTAATGATATATATGGGCA	5004	
OY	788	TACAAGCGCGGTTTGTATCAAGACCGGTAGCAGCAAAATATGAGGTAGTAG 838		
Db	5005	TATTAACCAACAGGTATATATAGGTATCAAGAGGTAAATGAAACAGTATCTGG 5055		
RESULT 15				
ACF69784				
ID	ACF69784 standard; DNA; 876 BP.			
AC	ACF69784;			
XX				
DT	20-NOV-2003 (first entry)			
XX				
DE	Phototriabodus luminescens nucleotide sequence #8251.			

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 PN MO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSD) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PI Duchaud B, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 2, SEQ ID NO 8251; 1205bp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 CC
 SQ Sequence 876 BP; 200 A; 177 C; 221 G; 278 T; 0 U; 0 Other;

Query Match 9.5%; Score 277.8; DB 7; Length 876;
 Best Local Similarity 62.2%; Pred. No. 3.2e-77;
 Matches 528; Conservative 0; Mismatches 252; Indels 69; Gaps 3;

QY 14 ATGCTCAGGGTTATGCAAGTTCCCAATATTGACCCCGTATGTTTCGATCGGCCT 73
 DB 1 ATGAGTAACAGCTACCTGCGCATTTCTTAATATTGACCCGTCATTTTCAATAGGTCCA 60
 QY 74 CTAGCGGTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 133
 DB 61 ATAGCCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
 QY 134 GCCAATCGCGGAGCGGATGCGGCGGAGTGTGAGCCGTGAGCAAGTCTCTGACTTG 193
 DB 121 GCGACACGAGGAGCGGAGCAACCAATATGTTGAGCAAAAAACAACTAGAGAACTTG 180
 QY 194 TTAATTCGCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 253
 DB 181 CTTTATGCAAGTTTGTGCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTAT 240

QY 254 AATTTGATCTGCTCTGCTGACCCCTCTTATTTTCAAGTGTGACGTGCGGACATG 313
 DB 241 AATTTCCGCTCTTCTGTGACAAACCCCTTATTTGTTAAAGTTTGGAGATGTGGATG 300
 QY 314 TCTTTCACGCGGCTTATGAGGTGATACCGGCATGTTCTGTATGCGGTAAAC 373
 DB 301 TCTTTCACGCGGCTTATGAGGTGATACCGGCATGTTCTGTATGCGGTAAAC 360
 QY 374 CAAGCCACTTCTTGTGAGCGGATTTGTTGCGCCCTTATGATGCAATGCGTTG 433
 DB 361 AAGCGCATTTCTTACAAAGTGTGATCTCATTTGCGCCGTATGCTCTTGTGATG 420
 QY 434 ATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 490
 DB 421 ATGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 QY 491 CCTGCGCTTTTGTATTCCTTA-----ATGATGCGCACGCGCGCATTCACAGCTT 512
 DB 481 CCGTGGCGATGTATTCACCAAGCTCCGATGTAAGATATTGATGATGATGATGATG 540
 QY 513 -----ATGATGCGCACGCGCGCATTCACAGCTT 547
 DB 541 CCTGACTGTATATCAGTACTGTAGCAATATGATGATGATGATGATGATGATGATG 600
 QY 548 TATGATTCGCTTATGAAAGCGGTGCTGTTCTTATTTCTTATTTATGTTATGTTAA 607
 DB 601 TATGAAATGCGCTTATGAAAGCATCTTCTGTTCAATTTATGATTTATATCCGTTAA 660
 QY 608 CCTGCTCGCTATGAGCAGGTATCGGACTGTTTTTATGCTGATACGTATTCGCTTC 667
 DB 661 CCGCCCAATGCGCAGATGTTCCGTTTATTTCTTATTTATGCTATGCTATGTTCCGAGTG 720
 QY 668 CTGTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 727
 DB 721 ATATGTCAGTCTTCTGCGCAACAGATGCACAACTGAGCTTATTTGACCGGATCAGT--- 777
 QY 728 ATGAGGCAAAATCTCTCTTACCTATGATGATGATGATGATGATGATGATGATGATG 787
 DB 778 ATGAGGCAAAATCTCTCTTATATGATGATGATGATGATGATGATGATGATGATG 837
 QY 788 TACAAGCGC 796
 DB 838 TATAAACAC 846

Search completed: July 31, 2004, 13:34:23
 Job time : 770 secs

This Page Blank (250.0)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 31, 2004, 12:39:50 ; Search time 150 Seconds
(without alignments)

10762.348 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909

Sequence: 1 gaggaagttgtatcgcctc.....cagcgaatcggcgctgcag 2909

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/pcodata/2/ina/5A COMB. seq: *
2: /cgn2_6/pcodata/2/ina/5B COMB. seq: *
3: /cgn2_6/pcodata/2/ina/5A COMB. seq: *
4: /cgn2_6/pcodata/2/ina/6E COMB. seq: *
5: /cgn2_6/pcodata/2/ina/6E COMB. seq: *
6: /cgn2_6/pcodata/2/ina/6E COMB. seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744.2	25.6	1830121	4 US-09-557-884-1	Sequence 1, Appl
2	744.2	25.6	1830121	4 US-09-643-990A-1	Sequence 1, Appl
3	410.8	14.1	873	4 US-09-543-681A-2282	Sequence 2282, Ap
4	292.2	10.0	807	4 US-09-377-565-17	Sequence 17, Appl
5	282.4	9.7	810	4 US-09-252-991A-15221	Sequence 15221, A
6	282.4	9.7	954	4 US-09-252-991A-15232	Sequence 15232, A
7	282.4	9.1	1434	4 US-09-543-681A-15227	Sequence 15227, A
8	265.8	8.8	879	4 US-09-489-039A-2365	Sequence 2365, Ap
9	256.8	7.2	834	4 US-09-328-352-3810	Sequence 3810, Ap
10	152.2	5.2	640681	4 US-09-790-988-1	Sequence 1, Appl
11	143.8	4.9	891	4 US-09-540-336-663	Sequence 663, App
12	142.2	4.9	96109	4 US-09-596-002-35	Sequence 35, Appl
13	115	4.0	435	4 US-09-252-991A-15244	Sequence 15244, A
14	75.4	2.6	2288	4 US-08-553-433-1	Sequence 1, Appl
15	75.4	2.6	2288	4 US-09-553-433-1	Sequence 1, Appl
16	75.4	2.6	4137	4 US-09-499-964-2	Sequence 2, Appl
17	75.4	2.6	2058	4 US-09-252-991A-15243	Sequence 15243, A
18	71.6	2.2	942	4 US-09-347-878-29	Sequence 29, Appl
19	64.8	2.2	1536	4 US-09-089-195-1	Sequence 1, Appl
20	64.8	2.2	1536	4 US-09-367-007C-38	Sequence 38, Appl
21	64.8	2.2	1536	4 US-09-962-665-6	Sequence 6, Appl
22	64.8	2.2	1536	4 US-09-963-333-6	Sequence 6, Appl
23	61.2	2.1	978	4 US-09-134-000C-941	Sequence 941, App
24	57	2.0	975	4 US-09-107-532A-2491	Sequence 2491, App
25	54.2	1.9	7685	4 US-09-221-017B-1092	Sequence 1092, App
26	49	1.7	969	4 US-09-134-001C-357	Sequence 357, App

28	48.2	1.7	870	4 US-09-252-991A-15224	Sequence 15224, A
29	47.4	1.6	801	4 US-09-489-039A-2884	Sequence 2884, Ap
30	46.6	1.6	867	4 US-09-107-532A-2355	Sequence 2355, Ap
31	46.6	1.6	1014	3 US-09-230-637-1	Sequence 1, Appl
32	46.6	1.6	35100	2 US-08-770-379-17	Sequence 17, Appl
33	46.6	1.6	35100	3 US-08-757-669A-17	Sequence 17, Appl
34	46.6	1.6	35100	4 US-09-230-637-1	Sequence 17, Appl
35	45.8	1.6	5924	4 US-08-956-171E-130	Sequence 130, App
36	45.4	1.6	2335	4 US-08-961-527-275	Sequence 275, App
37	44.8	1.5	852	4 US-09-134-001C-2654	Sequence 2654, App
38	44.8	1.5	870	4 US-09-134-000C-1317	Sequence 1317, App
39	42.6	1.5	621	4 US-08-935-433-3	Sequence 3, Appl
40	42.6	1.5	621	4 US-09-553-132-3	Sequence 3, Appl
41	42.4	1.5	1839	4 US-09-203-895-2	Sequence 2, Appl
42	40.6	1.4	7218	1 US-08-232-463-14	Sequence 14, Appl
43	40.6	1.4	1568	4 US-08-935-165A-95	Sequence 95, Appl
44	39.8	1.4	4214	4 US-09-221-017B-293	Sequence 293, App
45	38.6	1.3	7218	1 US-08-232-463-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 25.6%; Score 744.2; DB 4; Length 1830121;
Best Local Similarity 66.7%; Pred. No. 1.7e-245;
Matches 114; Conservative 0; Mismatches 538; Indels 27; Gaps 4;
QY 2 AGAAGTTGTATGCTCAGGAGTTATCTGCAGTTCCCAATATGACCCCGATATGTTT 61

Db	958086	ATAAGGAAATATATGATTCATATTTACTTCTTCCCACTTTGATCCGAGTATTTT	958145
Oy	62	TCGATCGC-----CTCTACGGGCGCGTGGTATGCGTTGATTTGGTGGTTTC	958115
Db	958146	AACGCTGGCGATGATGATATCGGCTTACCGTTGGTATGCGCTGATGATCACTTTAGGTTT	958205
Oy	116	CTTTTGTCTATGCGTTGGCCAAATCGCCGAGATCGCGCGGCGATGCGGCGT	958175
Db	958206	GTTTTTGACGTTGGCTGGCGGTTGGCCGTGCTAAATCGCCAAATGCGGTGGACAGTA	958265
Oy	176	GAGCAAGCTCTCGACCTTGTATTTCCGCGCGCTTTTATAGTATGATCGGTGGCCAGTT	958235
Db	958266	GATCAAGTTGATAGCTTACTTTCACACGGTTTTATGCGGGGTGTTATTTGGCGGACGTGT	958325
Oy	236	GGTATATGATCTTCTACAAATTTTGAATCTGTTCTTGTGACCCCTTATTTATTTCAAA	958395
Db	958326	GCGCAGTATTTTTTCTATATCTCGATCAATTTCTTACAGAAGACACTTATTTATTTCCG	958385
Oy	296	GTTGTCGATCGCGGCATGTCCTTCCACGCGCGCTTATGGGTGTATCACCGCATGTTTC	958355
Db	958386	GTTTGGGAAGGGGATATGTCGTTCCACGCGTGCCTTAATTTGGTATATTTGTTGTAATTT	958445
Oy	356	TGATATGGCGGTAATAAACGACGACCTCTTGTGTGTCGCGATTTGTGCGCCCTTAA	958415
Db	958446	TGACATCTTATTTCTCAAAAAGTATTTTGGCAACGCGTGAATTTGTGGCCTTTG	958505
Oy	416	GTGCAATTCGTTTGGGGATGGGAGTATCGGTAATTTATGATATGATCACTTTGGGGA	958475
Db	958506	ATTCGTTTGGTTTATGTTTATGAGGAAGATTTGGTAATTTCAATATCTTGAACATATGGGA	958565
Oy	476	CGAGTAAAGGATGTCCTTGGGCTTTGTATTCCTTAATGATGGCC--CACTGCGCGC	958532
Db	958566	CGCGAAACGAATGTCCTTGGGCAATGATTTTCCGAATGATCTCTTTTACGCTCGT	958625
Oy	533	CATCCTTACACCTTATGATTTGCGCTTAGAAGGGTGTCTGTTCTTATTTCTAAT	958592
Db	958626	CATTCATCAACATTTATGATAGCTTTTGAAGGCTGTGTGTTTTCGATTTCTGAAT	958685
Oy	593	TGGTTATTTGTTAAACCTGTCGCTAGGACAGCATTCGGAATGTTTTTATGCTGATAC	958652
Db	958686	ATTTTATTTAAABAAACACATGCTCAATGCTGCTGTTGACAGTTTATCTTAATTTGTTAT	958745
Oy	653	GGTACATTCGCTCTCTTGTGGAATACGTCGCTGACCCAGATGCTCATGTTGGTCTGTTT	958712
Db	958746	GCGCGTCTCGTTTATTTATGTTGGAATATGTCGTGAACCTGAAG-----TTGAAATATTC	958799
Oy	713	GGTGCCTCATTTCAATGGGGAATTCCTCTCTTAACATATGATGATCAATCGGTATTTTG	958772
Db	958800	TTTGGGATTTATTAACAGAGGAGGACGCTTTCCTTGCAGTATTTATGTTGTCCTTTC	958859
Oy	773	ATGATGTTTGGCTCTTACAAAGCGCGGTTGTATCAAGACCGGTAGCACAAATATGGGT	958832
Db	958860	ATTATGCGCTTGGGCTTATTCACGC-----AAAATGCGGTCAATAAATATGGAG	958907
Oy	833	AGTTAGTGAACAGATTTAGATCTTTGTACAGCGCATCGTCGATCAAGGTGTTGGGTT	958922
Db	958908	ATTTTATGAACATATCTTAGCTTGTGCCGCAATTTAGTGAAGGGGAATGGGTT	958967
Oy	893	GAAATGAAACGAACGGGCAAGCGTTGTTGACTGTGATTTAATGCCATTTGACTTACAT	959052
Db	958968	GCTATATGAACGTAAAGGTATGCGCTCACGCGTCAATTAATGCAATTTAGAAATATGAT	959027
Oy	953	GTGGGCAACATCAGTTTCTCTAGTGACTACACGCAAGTTTTGAAAGCTGCGCTA	959012
Db	959028	GTGGGCAATATCAATTTCCGCTGATTTACTACCCGTAAAGATTTATGGAAAGCGCGCAT	959087
Oy	1013	GCCGAGTTGCTCGCTATATTCGTGGTATCGAATATGCGCGGATTTTGGCAATTAGGT	959072
Db	959088	GCTGAATTTTATGTTATATTCGTGGATATGCAATATGCGCGCTGATTTCCGCGCACTTGGC	959147
Oy	1073	ACCAAACTGGGATGCTATATGCCAATTTAAACCAAGATGGCTCAACATCTTTACCGT	959132

```

Db      959148 ACGAAGAACTTGGGATGCTAATSCAGATGAAATGACAGCTTGCTTGCACAAATCCGACCTG 959207
QY      1133 AAAGGTGAGATGACATGAGGACCGCTGTATGCTGTTCAGGGTAGAGCTTGGGCTAGACCT 1192
Db      959208 AGAGCGCTGTAGATATGGGGCGCTATATGCTGTGCAGAGGACGACATGGCGTAAGCCT 959267
QY      1193 GATGTGGTGCATATTTGACCAAGTTGAAAAAAGATTTGTTGATGATTTTGAGCGGTGGCTGTAT 1252
Db      959268 AATGAGAAACTATATGATCAGCTACGTAAATTTGTTATATCTTAACGAAAGTATTTGAT 959327
QY      1253 GACCCAGGTGAATTTCTTAACCTTCTACATTCGCGGTGAATTTTACATGCGGCTGTTCGC 1312
Db      959328 GATAGAGGAGATTTTAACTTTTTTAACTTCGCGGGAATTTGATCTTGTTGTCCTCG 959387
QY      1313 CCTTGCAATGTACAGCCATCATTTTTCATTTGCTGTGGGGATACCTTGATCTCAACAGTACT 1372
Db      959388 CCTGTATGCAATACCCATACCTTTTCTCTTGCGGGGATACCTTTCATCTTACTAGCTAT 959447
QY      1373 CAGCGTTATGATGATGTGCGCTTGGGGTGGAATTTCAACATGATGTGACAGTTATGTGTC 1432
Db      959448 CAGCGTCTGTGATGTTCGCTTGGAATTTCAATTCAAATTCAGGTGTTACCTTC 959507
QY      1433 CTGCGCTGATGAGCACAGATCAACAGGAAAAAGCCGCGCTTGCGGTATCAAGATCGTC 1492
Db      959508 TTAGCACTTATGTCTCAATCAACGGGCAAAAAAGCCGGAAGCATATCATAAATTTGTG 959567
QY      1493 AATGCGCACTATTACCAAGATCACTCGAATTTGATGCGGATGTGACCTAAACGTGAG 1552
Db      959568 AATGCGCATATTATTAAGATCAGCTTGAGTTAATGCGGAAGGTCAACTTAAACCGCAG 959627
QY      1553 CCATGCCAGCGCCCTCAGTTCATATTCATCCAAAGATTAACACTGACAGATTTGGA 1612
Db      959628 CTTTCCCATTAACAAATTTAGAAATTAATCCAGATATTAACGCTTGAGATTTGGA 959687
QY      1613 ACTTGGTCACTTTGGATTTTGTGACGTCAACCGATATCAGTTCACAGATCTATTCA 1672
Db      959688 ACTTGGTCAAGATGATGATTTTAAAGTCGTGTGATCAATCCACAGAACCAATTTAA 959747
QY      1673 TACCGCTTTTCAGTCTAAT 1691
Db      959748 TATCCTTTTCCGCTCAAT 959766

RESULT 2
; Sequence 1, Application US/09643990A
; Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
            Mark D. Adams
            Owen White
            Hamilton O. Smith
            J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of
                    the Haemophilus influenzae Rd Genome, Fragments
                    Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,429
 FILING DATE: 1995-06-07
 APPLICATION NUMBER: 08/426,787
 FILING DATE: 1995-04-21
 ATTORNEY/AGENT INFORMATION:
 NAME: Kenley K. Hoover
 REGISTRATION NUMBER: 40,302
 REFERENCE/DOCKET NUMBER: PB186P1C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 301-610-5790
 TELEFAX: 310-309-8439
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1830121 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-643-990A-1

Query Match 25.6%; Score 744.2; DB 4; Length 1830121;
 Best Local Similarity 66.7%; Pred. No. 1.7e-245;
 Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY 2 AGAAGTTGTTATGCGGCTTATCGACAGTTCCCAATATGACCCCGATTGTT 61
 DB 958086 AATAGGAAATATGATTAATCAATATTTACTTCTCCCACTTTGATCGAGTATTTT 958145
 QY 62 TCGATCGG-----CCTTAGCGGTGCGCTGTGTATGCTGATGATTTGGGTTTC 115
 DB 958146 ACGCTTGGGATGATGATATCGGCTTACGTTGGTATGCTGATGATCTTTAGGTTT 958205
 QY 116 CTTTTCATATGTTGGTTCATCGCCAGCGGATCGCGGGGAGTGTGAGCGCT 175
 DB 958206 GTTTTTCACGTTGGCTTCGGCTTCGCGCTTAATCGCCAAATGCGGTTGACAGTA 958265
 QY 176 GAGCAAGTCTGACTTGTATTCGCGGCTTTTATGATGATGATGCGGTGCGAGTT 235
 DB 958266 GATCAAGTTGATGATCTTACTTTTCAACGGTTTATGCGGCTTTATGCGGAGCTGTT 958325
 QY 236 GGTATGATCTTCTACATTTTATGATCTGTTCTTCTGTCGACCCCTTTATTTATCA 295
 DB 958326 GCGAGTATTTTCTATATCTCGATCTTCTTCAAGAACCACTTTATTTATCCGC 958385
 QY 296 GTGTGACTGGGCGCATGTCCTTCCACGCGGCTTATGCGGTGATACCGCATGTC 355
 DB 958386 GTTTGGGAAGTGAATGTCGTTCCACGCTGCTTAATGTTGCTATGAT 958445
 QY 356 TGGTATGCGGTAAACCAACGACCTTCTTGTGTGCGCATTTTGTGCCCTTTA 415
 DB 958446 TGGACATCTTATCTCAAAAAGCTAATTTTGGCAACGCGTATTTTGTGGCCTTTG 958505
 QY 416 GTGCCATTCGTTTGGGATGAGACGTATCGGTACTTTATGAAATGTAACCTTTGGGA 475
 DB 958506 ATTCCTTTGTTTGGTTTGGCAGAAATTTGTAATTTCTTAATCTTGAATGGA 958565
 QY 476 CGAGTAAAGGATGTCCTTGGGCTTTGATTCCTTAATGTGGC---CACTGCGGCG 532
 DB 958566 CCGGAAACGAATGTGCTTGGGCAATGATTTCCGAATGATCTTTTACGCTCGT 958625
 QY 533 CATCTTCACAGCTTTATGAAATTCGCTTGAAGGGGTGTTCTGTTCTTTATCTTAAT 592
 DB 958626 CATCAATCAACTTATATAGCTTTTGAAGGCTGTGTGTTTACGATTTGAAAT 958685
 QY 593 TGGTTATATGTAACCTGTCGCTAGGACAGGTATCCGACGTGTTTATAGCTGATAC 652
 DB 958686 ATTTTATTAATAAACCAAGTCCAAATGCTTCTGTTGAGGTTTATCTTAATGTTAT 958745
 QY 653 GGTACATTCGCTTCTTGTGAATAGTCGCTGACGACGATGCTCAGTTGGTCTGTT 712
 DB 958746 GCGCTCTCGTTTATGTGTGAATATGTCGTAACCTGAAG-----TTGAAATATTC 958799

QY 713 GGTGCTTCATTTCAATGGGCAAAATCCTCTCTTACTATGATGATCATCGTATTTTG 772
 DB 958800 TTTGGATTAATACAGAGGCAAGCCCTTCTGCGATGATTAATGTTGTCCTTTC 958859
 QY 773 ATGATGTTTGGCTTTCAAGGCGGTTTGTATCAAGCCGTGTGACGCAAAATAGGCT 832
 DB 958860 AATATGCTTGGCTTATTCAGC-----AAAAGTCGTCAATAAATAGGAG 958907
 QY 833 AGTTAGTAAACAGATATTTAGATCTTTGTACGCGCATGTCGATCAAGGTTTGGCT 892
 DB 958908 ATTTTATAGACATATCTTGAAGCTTGTCCCGCATGTTATGTAAGAGGGAATGCTT 958967
 QY 893 GAAATGAACGAACGCGCAAGGCTTGTACTGTATTAATGCCATTTGACTACAT 952
 DB 958968 GCTAATGAACGTAAGGTAAGGATGCTTCAAGCTATTAATGAGATTTAGATATGAT 959027
 QY 953 GTGGCAACATCAGTTTCTTACTGACTACAGCAAGATTTTGGAAAGCTGCGTA 1012
 DB 959028 GTGGGAATTAATCAATTTCCGCTGATCTACCCGTAATAATGGAAGCCGCAAT 959087
 QY 1013 GCCAGTTGCTGGCTATATGCTGTTATGATTAATGCGGCGGATTTTCCCAATTAGT 1072
 DB 959088 GCTGAATTTTATGTTATATGTTGATATGCAATGCGCTGATTTTCCGCACTTGGC 959147
 QY 1073 ACCAAACCTGGATGCTAATGCAATTTAAACAAAGATGCTCAACATCTTACCT 1132
 DB 959148 ACGAAACCTGGATGCTAATGCAATTTAAATGCAAGCTTGGCTTCAAAATCCGATCT 959207
 QY 1133 AAGGTGAAGATGACATGGGACGCTGATGTTGTTCAAGGTAGAGCTTGGCTAAGCT 1192
 DB 959208 AAGGCGTTGATGATATGCGGCGCTATATGTTGTCAGAGCAAGGCAAGGTTAGCT 959267
 QY 1193 GATGATGCTATTTGACCAAGTTGAAAGATTTGTTGATTTTGAAGCCGTCGCTGAT 1252
 DB 959268 AATGAGAACTATGATAGCTACGTAAATTTGTAATTTAATTTAATCAAGATTTGAT 959327
 QY 1253 GACGAGGTGAATTTCTTAACTTCTACAAATCCGCGTGAATTTCAATGGGCTGTCGC 1312
 DB 959328 GATAGAGAGATTTTAACTTTTAACTTGGGGAATTTGATCTTGTGTTCTGCT 959387
 QY 1313 CCTTCATGATACGCAATCTTTTCAATTTGCTGGGCGATCTTGTATCTCAACAGTACT 1372
 DB 959388 CCTGTATGCAATACCAATCTTTTCTGTTGGCGATCTTATCATCTTACTAGCTAT 959447
 QY 1373 CAGGCTTACATGATGTCCTTGGGTTGATTTTCAACATGTCGACGTTTATGTTTC 1432
 DB 959448 CAGCGTCTCTGATGTTCCGCTGATTTGAATTTCAATCAAAATTCAGGTGTTTACCTTC 959507
 QY 1433 CTTCGCTGATGGCAGATCAAGGAAAGCCGCGCTTGGCTATCAACAATGCTC 1492
 DB 959508 TTAGCATTTATGCTACATCAAGGCAAAAGCCGCAAGCATATCATATAATTTGTG 959567
 QY 1493 AATGCGCATTTTACCAAGATCAATCGAATTTGATGCGGATGTCAGCTAAACGTGAG 1552
 DB 959568 AATGCGCATTTTATGAAGATCAGCTTGAATTTGCGTGAATCACTTAAACGCGAG 959627
 QY 1553 CCATTCACAGCCCTCAGTTCATATCAATCAAAAGATTTAAACCTGAGATTTTGA 1612
 DB 959628 CTTTCCCATTTCAAAATTTAATTAATCAAGATTTAAACCTTGAATTTTGA 959687
 QY 1613 ACTTGGGTACTTTGATGATTTTGAACGTCACGGAATACGTTTCAAGATCTTATCA 1672
 DB 959688 ACTTGGGTCAATGATTTTAAAGTGTGCTATCATCTCCAGAACCAATTA 959747
 QY 1673 TACCGTTTTCAGTCAAT 1691
 DB 959748 TATCTTTTTCGCTCAAT 959766

RESULT 3
 US-09-543-681A-2282
 ; Sequence 2282, Application US/09543681A
 ; Patent No. 6605709

```
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/09/543,681A
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 2282
/ LENGTH: 873
/ TYPE: DNA
/ ORGANISM: Proteus mirabilis
/ US-09-543-681A-2282

Query Match      14.1%; Score 410.8; DB 4; Length 873;
Best Local Similarity 67.0%; Pred. No. 5,2e-132;
Matches 583; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

QY 820 AGCAAAATAGGAGTACCTAGTAAACAGTATTAGTCTTTGTGACGGCATGTCATCA 879
Db 3 AACAATAAAGAGAGACTACATAGAGAGATCTGGCATTTGTCAACCATTAATGATGA 62
QY 880 AGGTGTTGGGTTGAAATGAAAGCAAGCGGCAAGCGTTGTTGACTGTGATTAAATGCCGA 939
Db 63 AGGTAAATGATGATTAAATTAATTAACGACAGAACCGGTTGTTAAACGGTGAATAAGCCGA 122
QY 940 TTTGACCTACATGTTGGGCAACATCAATGTTTCTCTAGTACTACACGAGAGTTTGG 999
Db 123 TCTGGAATATGATGTTGGCAATTAACCAATTCACCTACACCAACGCGTAAAGTTTGA 182
QY 1000 GAAACCTGCGGAGCGAGTGTCTGCGCATATGTTGGTGTTCAGTAAATGCGCGGATTT 1059
Db 183 CAACGCGGATGTTGGGAACTTAATAGGTTACTTACGTGATTAATGACAGCAATTT 242
QY 1060 TCGCAATTTAGGTACCAAACTGGGATGCTAATGCCAATTTAAACCAAGATGGCTCA 1119
Db 243 TCGTCTATTTGGCTGTAATACGTGAACGCAATGCTAATGAATAATAGCGATGGTTAA 302
QY 1120 CAATCTTAACGTTAAAGTTGAGATGACATGGGAGCGGTGATGTTGTTCAAGGTTAGAC 1179
Db 303 TAATCCGATCGTAAGGTTGAGATGATATGAGCGGAGCTATGTTGTTCAAGGACGCA 362
QY 1180 TTGGGCTAAGCGTAAAGTTGAGATGACATGGGAGCGGTGATGTTGTTGAG 1239
Db 363 ATGGCAAGCGCTGATGCTGCGCATTTGATCAATTAAGTAAATGTTGATTAATCC 422
QY 1240 CCGTGGCGTTGATGACCGAGTGAATTTCTTAATCTTACATCCGGGTGAATTTCAAT 1299
Db 423 TAACGATATGACGATCGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 1300 GGGGTGTTTGGCGCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
Db 483 AGGGTGTATTAACGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
QY 1360 TCTCAACAGTATGACCGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 1419
Db 543 TTTAATCTCAATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
QY 1420 GGTATATGTTTCTTGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
Db 603 ATGCTTTGATTAATTAAGCGCTGTGCTCAATTCACAGGCGATTAACAGGTAAGCATTT 662
QY 1480 TCACAAGATCGTCAATGCGCATATTACCAAGATCACTGAAATGATGATGATGATGATGAT 1539
Db 663 CCATTAAGATTAATTAATGCTCAATTAATGAACAATTAATCTGATGATGATGATGATGAT 722
QY 1540 GGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
Db 723 ACTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
QY 1600 GAGGATTTGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
```

```
Db 783 AGATGATATAGAACTTGGGTACAGACAGATGATGATGATGATGATGATGATGATGATGAT 842
QY 1660 CGATCTTATCAATACCGGTTTCAGTCTA 1689
Db 843 CGAGCGATAAATTAACCGTTCAACGGTGA 872

RESULT 4
US-09-277-565-17
/ Sequence 17, Application US/09277565
/ Patent No. 6207384
/ GENERAL INFORMATION:
/ APPLICANT: Mekalanos, John J.
/ APPLICANT: Akereley, Brian J.
/ APPLICANT: Rubin, Eric J.
/ APPLICANT: Camilli, Andrew
/ TITLE OF INVENTION: SYSTEMATIC IDENTIFICATION OF ESSENTIAL
/ TITLE OF INVENTION: GENES BY IN VITRO TRANSPOSON MUTAGENESIS
/ FILE REFERENCE: 00742/052002
/ CURRENT APPLICATION NUMBER: US/09/277,565
/ EARLIER FILING DATE: 1999-03-26
/ EARLIER APPLICATION NUMBER: 60/079,770
/ NUMBER OF SEQ ID NOS: 29
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 17
/ LENGTH: 807
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae
/ US-09-277-565-17

Query Match      10.0%; Score 292.2; DB 3; Length 807;
Best Local Similarity 65.0%; Pred. No. 1e-90;
Matches 467; Conservative 0; Mismatches 243; Indels 9; Gaps 2;

QY 81 TCGGTGTATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 140
Db 74 TACGTGTATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 133
QY 141 GCCAGCGGATGCGCGGCGGAGTGTGAGACCGGTGAGCAAGTCTGATCTGATGATGATGAT 200
Db 134 GCGGCTTAATGCGGCAATATGCGGTGAGACAGATGATGATGATGATGATGATGATGATGAT 193
QY 201 CCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 260
Db 194 ACGGTTTATGAGGCTGTTTATGAGCGAGCTGTTGCGATGATGATGATGATGATGATGATGAT 253
QY 261 ATCTGTTCTTGTGACCGCTTATTAATTAATCAAGTGTGATGATGATGATGATGATGATGAT 320
Db 254 ATCAATTTCTCAAGAACCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 313
QY 321 ACGGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 380
Db 314 ACGGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 373
QY 381 CTTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 440
Db 374 ATTTTGGCAAAAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 433
QY 441 GTATCGGTAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 500
Db 434 GAATTTGTAATTTTATTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAAT 493
QY 501 TTTGATTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
Db 494 TGATTTTCCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553
QY 558 CTTTAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
Db 554 TTTTAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 613
QY 618 TAGGAGGATATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 677
```


Db 614 TGGCTTCTGTCAGAGTTTATCTTATTTGTTATGTCGCTTCCCTTTTATTTGTAAT 673
Qy 678 AGCTGCTGAGCCAGATGCTGAGTTGGGCTGTTTGGGCTTCATTTCAATGGGGCAAA 737
Db 674 ATGTGCTGTAACCTGAAG-----TTGAAAATTTCTTTGGGATTTATTAACGAGGGCAAG 727
Qy 738 TCCCTCTCTTACCTATGATGATCATCGGTAATTTGATGATGTTGGTCTTACACGCG 796
Db 728 CCTTGTGCTGGCGATGATTTATGATGCTTCTTATTTAGCTTGGGCTTATTCAGCG 786

RESULT 5
US-09-252-991A-15221
; Sequence 15221, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15221
; LENGTH: 810
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15221

Query Match 9.7%; Score 282.4; DB 4; Length 810;
Best Local Similarity 61.3%; Pred. No. 2,6e-87;
Matches 473; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

Qy 29 CTGCAATTTCCCAATATTGACCCCGATTTGTTTGGATGCGCCCTCTAGCGGCTGCG 88
Db 13 CTGAGTATCCCAATGATCATCGGTTGCGCTGCGCATGCGCCGCTGAAGATCACTG 72
Qy 89 TATGCTTGATGATTTGGTGGGTTTCTTTTGTATGTTGTTGCGCAATGCGCGAGCG 148
Db 73 TACGGGCTGATGATCTGATCGGAGATCGGCGCGCTGCTGCTGCGCTGCGCGAG 132
Qy 149 GATCGCGCGGAGATGTTGAGACGCTGACGCAAGTCTCTGACTTTTATTCGCGGCTTT 208
Db 133 AAGCGCTTGACCCGACCTGACCAAGAGCGCTTCCGACCTGCTTCTGCGTCCG 192
Qy 209 TTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
Db 193 TCGGGGATGATCTCGGTGCGCGCTGCGCTGATGCTGCTTCTACAACTGACGAGTAC 252
Qy 269 CTGCTGACCTCTTTATTTATTTCAAGTGTGATGCGGAGATGCTTTCACGCGGCG 328
Db 253 ATGCGCAACCCAGCCTGATCTTCAAGTCTGGAAGGCGGAGATGCTTTCATGCGGCG 312
Qy 329 TTAATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 388
Db 313 CTGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 372
Qy 389 GGTGTCGCGGATTTTGTTCCTTTAGTCCATTTGTTGGGATGGAAGTATCGGT 448
Db 373 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 432
Qy 449 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 508
Db 433 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 492
Qy 509 CCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
Db 493 CCAAGGCGGCGCGCTGCGCGGAGATCTCTGACACTGATCCAGTTCCCGCTGGAAGGC 552

Qy 569 GTGTTCTGTTCTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 628
Db 553 GTGGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 612
Qy 629 TCCGAGCTGTTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
Db 613 TCCGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 672
Qy 689 CAGATGCTCAGTTGGG---TCTGTTGTTGCTTCAATTTGATGAGGCAAACTCTCC 745
Db 673 CCGATGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
Qy 746 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 797
Db 733 GTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784

RESULT 6
US-09-252-991A-15232/C
; Sequence 15232, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15232
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15232

Query Match 9.7%; Score 282.4; DB 4; Length 954;
Best Local Similarity 61.3%; Pred. No. 3e-87;
Matches 473; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

Qy 29 CTGCAATTTCCCAATATTGACCCCGATTTGTTTGGATGCGCCCTCTAGCGGCTGCG 88
Db 945 CTGAGTATCCCAATGATCATCGGTTGCGCTGCGCATGCGCCGCTGAAGATCACTG 886
Qy 89 TATGCTTGATGATTTGGTGGGTTTCTTTTGTATGTTGTTGCGCAATGCGCGAGCG 148
Db 885 TACGGGCTGATGATCTGATCGGAGATCGGCGCGCTGCTGCTGCGCTGCGCGAG 826
Qy 149 GATCGCGCGGAGATGTTGAGACGCGTGAAGTCTCTGACTTTTATTCGCGGCTTT 208
Db 825 AAGCGCTTGACCCGACCTGACCAAGAGCGCTTTCGACCTGCTGCTGCGGTCCG 766
Qy 209 TTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
Db 765 TCGGGGATGATCTCGGTGCGCGCTGCGCTGATGCTGCTTCTACAACTGACGAGTAC 706
Qy 269 CTGCTGACCTCTTTATTTATTTCAAGTGTGATGCGGAGATGCTTTCACGCGGCG 328
Db 705 ATGCGCAACCCAGCCTGATCTTCAAGTCTGGAAGGCGGAGATGCTTTCATGCGGCG 646
Qy 329 TTAATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 388
Db 645 CTGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 586
Qy 389 GGTGTCGCGGATTTTGTTCCTTTAGTCCATTTGTTGGGATGGAAGTATCGGT 448
Db 585 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 526
Qy 449 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 508

QY 326 GGCCTATTGGGTGTGATCAACCCGATGTTCTGGTATGCGCGTAAACCAACGACCTTC 385
 DB 319 GGCCTATTGGGTGTGATCAACCCGATGTTCTGGTATGCGCGTAAACCAACGACCTTC 378
 QY 386 TTGCTGTGCGCGATTTTGTGTCCTTTAGTGCATTCGTTTGGGATGAGAGATATC 445
 DB 379 TTTCAGATGCGACACTTGTGTCATTAATCCCATTTGGTTTGGTTAGGTTAGGACGTATT 438
 QY 446 GGTACTTATTAATAGTACCTTGGGGAAGATAC---GGATGTGCTTGGGCTTTT 502
 DB 439 GGTAACTTATTAATAGTACCTTGGGGAAGATACCTTGGGGAAGATACCTTGGGGAAT 498
 QY 503 GATTTCCC----- 510
 DB 499 TTATTTCCAGGCTCCCGTGCAGAAATCTACAGCTTGTGCGCAAGATCCACTACACTA 558
 QY 511 -----TAATGTGCGCCCACTGCGCGCATCTTCACAGCTTTATGAATTC 556
 DB 559 TTACCTATCATTCAGCAATACGCGATATGCTCGCCATCCGCACTCTATGAATG 618
 QY 557 GCTTATGAAGCGGTGTTCTTCTTATTTCTTAATGTTATGTTATGTTAACTGTCGG 616
 DB 619 TTCTTAAGAGGCGGTGTTATTTATTTCTGAATATCTTGTGCGCAAAATCGTCT 678
 QY 617 CTAGCAGCGTATCCGAGCTGTTTATGCTGATACGGTACATCCGCTTCTGTTGGA 676
 DB 679 ATGGGAGATGATCGGCTTATTTCTTAATGTTATGTTGCTTCCGATTAATGTTGA 738
 QY 677 TACGTCGTGACAGATGCTGATGTTGGTGTGTTGTTGCTTCAATGAGGCGAA 736
 DB 729 TTCTTCGCGCACTGATGCGCAATTAAGTTTATTAAGCGTATCA---GCATGGGCGAA 795
 QY 737 ATCTCTCTTACCTATGATGATCATCGGTATTTGATATGTTGTTGTTTACAAAGCC 796
 DB 736 ATCTTTCATTCATGATGATTAATTAAGCATTAATCTTATGATGTTGCGCGATCGTCC 855
 QY 797 GGT 799
 DB 856 GAT 858

RESULT 9 US-09-489-039A-2965

; Sequence 2965; Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; PRIORITY FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIORITY FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 2965
 ; LENGTH: 879
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-2965

Query Match 8.8%; Score 256.8; DB 4; Length 879;
 Best Local Similarity 59.5%; Pred. No. 2.3e-78;
 Matches 515; Conservative 0; Mismatches 282; Indels 69; Gaps 2;

QY 14 ATGCCGAGGGTTATCTGAGTTCCCATATGACCCGATGTTTTCGATCGGCCCT 73
 DB 4 ATGAATAGTGGCTACCTGATTTTCCGAGTTGATCGGATATTTTCTCAGGCCG 63
 QY 74 CTAGCGGTGCGGTGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 133
 DB 64 GTCCTCCTCACTGATGATGTTGATGATGATGATGATGATGATGATGATGATGATG 123

QY 134 GCCATCCCGAGCGGATGCGCGGAGTGTGAGCGCGTGAAGATCTGACTTG 193
 DB 124 GCCATCCCGAGCGGATGCGCGGAGTGTGAGCGCGTGAAGATGAGTGAAGTGAAGTGA 183
 QY 194 TTATTCGCGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 253
 DB 184 CTCTAGCGGCGGCTTCTGCGGCTTCTTCCGCGCGCGCTATCGGCTAGGCTGTTTAC 243
 QY 254 AATTTGATCTGTTCTGCTGACCCCTTATTAATTAAGTGTGATGAGTGGCGG 313
 DB 244 AACCTGCGGATTTCTTGCAGATCCGCTGATCTGTTCCGCTGTGAGCGGCGG 303
 QY 314 TCCTTCAAGCGGCTTATGAGTGTGATCAACCGCATGTTCTGATGCGGTTAAAC 373
 DB 304 TCGTTCAAGCGGCTTATGAGTGTGATCAACCGCATGTTCTGATGCGGTTAAAC 363
 QY 374 CAAGCACCTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 433
 DB 364 AAACGACCTTCTTCAAGTCTCGATTTTATGCGCGCTGATCCGTTGGTCTTGGC 423
 QY 424 ATGGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 481
 DB 424 GCGGCGGCTGAGGCACTTATCAACGCGAGCTGTGAGCGCGGCTGAGACCGAGCTTC 483
 QY 482 -----ACGAT 487
 DB 484 CACTACAGATGATTTTCCCGGCTCCCGTGCAGAACCTGCGCTGCTGCGGACGAT 543
 QY 488 GTGCTTGGGCTTTTGTATTCCTAATGATGAGCCCACTGCGCGCATCTTACAGCTT 547
 DB 544 CCGAATGAGCAATGCTGTTGATACCTACGCGCGGCTGCGCGGCTGAGCGG 603
 QY 548 TATGATTCGCTTATGAAGGCGTGTGTTCTTATTTCTTATTTGTTATTTGTTAA 607
 DB 604 TACGAGCTGCGCGTGAAGGCGTGTGATCTGTTCTGATCTCAACCTGTTATTCGTTAA 663
 QY 608 CCGTGTGCGGCTGAGGAGGATTCGCGGCTGTTTATGCTGATGAGTACGATTCGCTTC 667
 DB 664 CCGGCGCGGAGCGGCTCGTTGCTGTTGCTGATGATGATGATGATGATGATGATG 723
 QY 668 CTGTGATATAGTCTGCGGAGGATGCTGATGTT---GGTCTGTTGTTGCTTCAAT 724
 DB 724 ATCGTGAATTTCTTCCGACGAGCCGACGATTCACCGGCGGTGAGTGAATAC 783
 QY 725 TCAATGCGGCAATCTCTCTTACCTATGATGATGATGATGATGATGATGATG 784
 DB 784 AGCATGCGGAGATTTCTTCAATCCGATGATGATGATGATGATGATGATGATG 843
 QY 785 TCTTCAAGCGGCTTGTATCAAGA 810
 DB 844 GCTTACCGGCAACCGTCCGAGCCACA 869

RESULT 10

US-09-328-352-3810
 ; Sequence 3810; Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-039A
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; PRIORITY FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 3810
 ; LENGTH: 834
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-3810

Query Match 7.2%; Score 208.6; DB 4; Length 834;

APPLICANT: Gary L. Breton et al.

RESULT 11-988-1
 US-09-790-988-1
 ? Sequence 1, Application US/09790988
 ? Patent No. 6632935
 ? GENERAL INFORMATION:
 ? APPLICANT: SHIGENOBU, SHUJI
 ? APPLICANT: MATSANABE, HIDEMI
 ? APPLICANT: HATTORI, MASHITA
 ? APPLICANT: SAKAKI, YOSHIYUKI
 ? TITLE OF INVENTION: GENEOME DNA OF BACTERIAL SYMBIONT OF APHIDS
 ? FILE REFERENCE: 081356/0159
 ? CURRENT APPLICATION NUMBER: US/09/790,988
 ? CURRENT FILING DATE: 2001-02-23
 ? PRIOR APPLICATION NUMBER: JP2000-107160
 ? PRIOR FILING DATE: 2000-04-07
 ? NUMBER OF SEQ ID NOS: 7
 ? SOFTWARE: PatentIn Ver. 2.1

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; LENGTH: 891
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-663

Query Match
Best Local Similarity 59.4%; Score 143.8; DB 4; Length 891;
Matches 263; Conservative 0; Mismatches 177; Indels 3; Gaps 1;

QY 38 CCCAATATTGACCCCGTATTGTTTGCATGCGCCCTTAGCGGTGCGCTGATGCTTG 97
DB 28 CCACAGATTAACTCTGCTTGTAGACTTGGCGATGATGAGCTGATGCTGATGCTTG 87
QY 98 ATGATTTGGTGGCTTCTTTTGTATGATGCTTGGCCATGCGCGATGCGGATCGCGG 157
DB 88 ATGATTTATTTGCTTTTGTGAGTGGCTTATTTTGGCATGGGTGCAACAA--ATCC 144
QY 158 GGCAGTGTGAGCGCGTGAAGCAAGTCTGACTGATTTATTCGCCGCTTTTATGCTGA 217
DB 145 CGAAGCATTTTAACTCTGATGATGCTTGGATTTGTTTGGTTCAGTGGGTG 204
QY 218 GTGATGCGTGGCGGATGCTGATGATGCTTCTTACAAATTTGATCTGTTCTGCTGAC 277
DB 205 ATTTTATGTCGCGTATTTGCTATGATCTTTTATACAAATTTGATGCTGATGCTAAC 264
QY 278 CCTCTTATTATTAACAAGTGTGAGTGGCGGATGCTCTTCCAGCGCGCTTATGGGT 337
DB 265 CCTTATATCTTTTTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 324
QY 338 GTGATACCGCATGTTCTGTATGCGCGTAAACCAACGACCTTCTTGGTGTGCGC 397
DB 325 GTGTCCTTCCATGCTGATTTTGGACATTAATATATATATATATATATATATATATAT 384
QY 398 GATTTTGTGCGCCCTTATGATGCGCATTCGTTGGGAGTGGAGCGTATCGTAACTTATG 457
DB 385 GATTTTATGCAACCTGCTGACCAACAGGCTTATTTTGGGGAATTTGGGAATTTTATC 444
QY 458 AATAGTGAACCTTTGGGAGCAAGT 480
DB 445 AATGGGAGCTTTGGGAGCAAGT 467

RESULT 13
US-09-596-002-35
; Sequence 35, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Paterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: Incyte template ID No. 6632636 35
; PUBLICATION INFORMATION:

```

```

US-09-596-002-35

Query Match
Best Local Similarity 59.1%; Score 142.2; DB 4; Length 96109;
Matches 262; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 38 CCCAATATTGACCCCGTATTGTTTGCATGCGCCCTTAGCGGTGCGCTGATGCTTG 97
DB 51130 CCACAGATTAACTCTGCTTGTAGACTTGGCGATGATGAGCTGATGCTGATGCTTG 51189
QY 98 ATGATTTGGTGGCTTCTTTTGTATGATGCTTGGCCATGCGCGATGCGGATCGCGG 157
DB 51190 ATGATTTATTTGCTTTTGTGAGTGGCTTATTTTGGCATGGGTGCAACAA--ATCC 51246
QY 158 GGCAGTGTGAGCGCGTGAAGCAAGTCTGACTGATTTATTCGCCGCTTTTATGCTGA 217
DB 51247 CCACAGATTAACTCTGAGATGCTTGGATTTGTTTGGTTCAGTGGGTG 51306
QY 218 GTGATGCGTGGCGGATGCTGATGATGCTTCTTACAAATTTGATCTGTTCTGCTGAC 277
DB 51307 ATTTATGTCGCGGATTTGCTATGATGCTTATTAACAATTTGATGCTGATGCTAAC 51366
QY 278 CCTCTTATTATTAACAAGTGTGAGTGGCGGATGCTCTTCCAGCGCGCTTATGGGT 337
DB 51367 CCTTATATCTTTTTCAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 51426
QY 338 GTGATACCGCATGTTCTGTATGCGGCTAAACCAACGACCTTCTTGGTGTGCGC 397
DB 51427 GTGTCCTTGCATGCTGATTTTGCATATTAATATATATATATATATATATATATAT 51486
QY 398 GATTTTGTGCGCCCTTATGATGCGCATTCGTTGGGATGAGGAGTATCGTAACTTATG 457
DB 51487 GATTTTATGCAACCTGCTGACCAACAGGCTTATTTTGGGGAATTTGGGAATTTTATC 51546
QY 458 AATAGTGAACCTTTGGGAGCAAGT 480
DB 51547 AATGGGAGCTTTGGGAGCAAGT 51569

RESULT 14
US-09-252-991A-15244/c
; Sequence 15244, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rudenfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15244
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15244

Query Match
Best Local Similarity 59.2%; Score 115; DB 4; Length 435;
Matches 196; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 29 CTGAGTTTCCCAATATTGACCCCGTATTGTTTGCATGCGCCCTTAGCGGTGCGCTTG 88
DB 373 CTGACGTATCCCAATGATGATCGGTGCGCTGCGCATGCGCCCTGAATATCCACTGG 314
QY 89 TATGCTTGATGATTTGATGCTTCTTTTGTATGCTGATGCTGATGCTGATGCTGATGCTG 148
DB 313 TACGGGCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 254

```

QY 149 GATCGCGGGGAGTGGTTGACGCGTGAACAAGTCTCTGACTGTATATGCGCGCTTT 208
DB 253 AAGCGCTTGAACCCGACCTGACCAAGAGCGCCTTTCGACCTGCTTCTGCGGTCCGC 194
QY 209 TTAGGTGAGTATGAGTGGGCGGAGTGGTATGATGATCTTAATAATTTGATCTGTT 268
DB 193 TCGCGGATGATCTGAGTGGCGCGCTGCGCTAGTGTCTTCAACCTGAGCAGATAC 134
QY 269 CTTCGACCCCTCTTATTTATTTCAAGTGTGACTGGCGGAGTCCCTTCCACGCGGAC 328
DB 133 ATCGCAACCCGACCGCTATCTTCGAGGCTGTGAAAGGCGCATGCTCTTCATGCGCGC 74
QY 329 TTATTGGGTGATCAACGCCCATGTTCTGCT 359
DB 73 CTGCTGCGGTATGCTGCGGCTGTGCTGCT 43

RESULT 15

US-08-935-433-1/c
; Sequence 1, Application US/08935433
; Patent No. 6319688
; GENERAL INFORMATION:
; APPLICANT: FEILD, JOHN
; TITLE OF INVENTION: A HUMAN SODIUM DEPENDENT PHOSPHATE
; TITLE OF INVENTION: TRANSPORTER (1PT-1)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,433
; FILING DATE: 23-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/044,974
; FILING DATE: 28-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-935-433-1

Query Match 2.6%; Score 75.4; DB 4; Length 2288;
Best Local Similarity 56.0%; Pred. No. 7.8e-15;
Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGGTAGAGATGACACCGCAGAAAGATGAAAAATCGTGGATGATGAC 2535
DB 771 CACCTTCACGGGCAAGAGACCAACGACGACGACGACGACGACGACGACGACGACG 712
QY 2536 ACTGCGAAGACGCGGCAAACTCTTCTTACAGCGCATATGCGCAAGGCTGACGAGAT 2595
DB 711 TCTCTCAAAAGCTTCTGAACTCACTCCGATCTCCCACTGATGAGCGCAACAATAGT 652

QY 2596 ATTGATCAAGTATGATACCAATATTTGGACCCCATCAACATAGAAATCGCGTTTCAACCG 2654
DB 651 GTTGGATGATGACGTTCCAAATGTTGGCCCCCATATATATGGGATGCGACGCCCAACAGT 592
QY 2655 --GTAAACCAACGGCAAGAGACCAACAAATATGAAGTCAACGCTGTTGAGATTGAAT 2712
DB 591 GAGCAATGAAAGAGACACCATGCTGACCAACGATGAGCGTTGAGGTGCTGAGCTCTGCAC 532
QY 2713 CAGTCCGTTGCCCACTAAACCAATCATCATCTGCAATTGGGTGGAGCA 2765
DB 531 CAAAGGTTCAACAGCACCCCGATCACAGGCCCAACAAAGGTTTGACATTA 479

Search completed: July 31, 2004, 17:02:39
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 15:39:05 ; Search time 888 Seconds
(without alignments)
16062.201 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909

Sequence: 1 gagaagcttgatgcctc.....cagcaaatcgagctgcag 2909

Scoring table: IDENTITY_NUC
Gapop: 10.0, Gapext: 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : Published Applications NA.*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1112.6	38.2	1149	US-10-282-122A-40937	Sequence 40937, A
2	842.4	29.0	852	US-10-282-122A-40936	Sequence 40936, A
3	744.2	25.6	1830121	US-10-329-960-1	Sequence 1, Appli
4	744.2	17.4	852	US-10-282-122A-30692	Sequence 30692, A
5	505.4	15.9	852	US-10-282-122A-22162	Sequence 22162, A
6	462.2	13.9	849	US-10-282-122A-12325	Sequence 12325, A
7	403.8	8.5	876	US-09-741-669-182	Sequence 182, App
8	248.2	5.2	640681	US-09-790-988-1	Sequence 1, Appli
9	152.2	4.9	96109	US-10-672-787-35	Sequence 35, Appli
10	142.2	4.0	2731748	US-10-297-465A-1	Sequence 1, Appli
11	115.4	2.9	2440	US-10-152-319A-1589	Sequence 1589, Ap
12	84	2.7	3950	US-10-152-319A-1988	Sequence 1988, Ap
13	77.8	2.6	2208	US-10-276-774-1331	Sequence 1331, Ap
14	75.4				

15	75.4	2.6	2280	US-10-173-999-149	Sequence 149, App
16	75.4	2.6	3952	US-10-116-802-36	Sequence 36, Appli
17	75.4	2.6	4137	US-10-052-664-2	Sequence 2, Appli
18	73.8	2.5	4111	US-10-257-021-146	Sequence 146, App
19	73.8	2.5	4135	US-10-097-320-220	Sequence 220, App
20	72.2	2.5	687	US-10-404-460-258	Sequence 258, App
21	70.2	2.4	1800	US-10-169-395-71	Sequence 71, Appli
22	70.2	2.4	2016	US-10-169-395-81	Sequence 81, Appli
23	70.2	2.4	2124	US-10-108-260A-1011	Sequence 1011, Ap
24	70	2.4	1164	US-09-738-626-3017	Sequence 3017, Ap
25	70	2.4	1333	US-10-627-476-511	Sequence 511, App
26	70	2.4	3309400	US-09-738-626-1	Sequence 1, Appli
27	69.2	2.4	969	US-10-282-122A-33224	Sequence 33224, A
28	68.2	2.3	795	US-10-282-122A-42178	Sequence 42178, A
29	67.6	2.3	969	US-10-282-122A-14737	Sequence 14737, A
30	66.4	2.3	792	US-10-282-122A-12730	Sequence 12730, A
31	65.4	2.2	722	US-10-332-859-254	Sequence 254, App
32	64.8	2.2	1478	US-10-084-817-332	Sequence 332, App
33	64.8	2.2	1536	US-09-877-178-11	Sequence 11, Appli
34	64.8	2.2	1536	US-10-342-887-556	Sequence 556, App
35	64.8	2.2	1536	US-10-172-118-556	Sequence 556, App
36	64.8	2.2	1536	US-10-388-360-350	Sequence 350, App
37	64.8	2.2	1536	US-10-159-563-244	Sequence 244, App
38	64.8	2.2	1536	US-10-283-975A-254	Sequence 254, App
39	64.8	2.2	3298	US-10-240-965-200	Sequence 200, App
40	64.6	2.2	795	US-09-741-669-181	Sequence 181, App
41	64.6	2.2	795	US-10-282-122A-6574	Sequence 6574, Ap
42	64.6	2.2	798	US-09-738-626-942	Sequence 942, App
43	64.6	2.2	1200	US-09-954-197-1	Sequence 1, Appli
44	64.6	2.2	2730	US-09-987-446-1	Sequence 1, Appli
45	64.4	2.2	969	US-10-282-122A-14234	Sequence 14234, A

ALIGNMENTS

```
RESULT 1
US-10-282-122A-40937/c
Sequence 40937, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zykkind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
```



```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40937
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Vibrio cholerae
US-10-282-122A-40937
```

```
Query Match      38.2%; Score 1112.6; DB 13; Length 1149;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1118; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1783 TTAAAGCGTTAGGGCAAGATGCTGCGGGGATGACGACAAACACCCAAATTAAGTACT 1842
DB 1149 TTAAAGCGTTAGGGCAAGATGCTGCGGGGATGACGACAAACACCCAAATTAAGTACT 1090
QY 1843 CACCAACCATTTGCTCTTAACAAGCCCAAGTTGATGAGTCACTTTAATAGG 1902
DB 1089 CACCAACCATTTGCTCTTAACAAGCCCAAGTTGATGAGTCACTTTAATAGG 1030
QY 1903 CAGTTGCGCTAAGAAAGAAATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 1962
DB 1029 CAGTTGCGCTAAGAAAGAAATACCGTAATCAAGACCGTAGCCATCAAGTTAAAGCTTAA 970
QY 1963 GTGCAACGCGCAATTTGCAAGCAAAACGCGCAAACTCAACAAGACACGCGTTGGCGG 2022
DB 969 GTGCAACGCGCAATTTGCAAGCAAAACGCGCAAACTCAACAAGACACGCGTTGGCGG 910
QY 2023 GAGCAGAGAGTAATCAAGTGCCTAATGTCGACCTAAGTAATGGTGAATTTTCACG 2082
DB 909 GAGCAGAGAGTAATCAAGTGCCTAATGTCGACCTAAGTAATGGTGAATTTTCACG 850
QY 2083 CACTTTGACGACGCGACGAGCCGACGAGAAACCAATTAGCTGTTGTCGATGAAGA 2142
DB 849 CACTTTGACGACGCGACGAGCCGACGAGAAACCAATTAGCTGTTGTCGATGAAGA 790
QY 2143 TTGAACATAATCCGTACCACTGTACCTGAAGCAATACCGTGTAGTGGCCCTCGCCAT 2202
DB 789 TTGAACATAATCCGTACCACTGTACCTGAAGCAATACCGTGTAGTGGCCCTCGCCAT 730
QY 2203 CGCATTTTGTGAATTTCAAGTGGCGGCAACCAATCAACTCTTCATCAATTTGCCAT 2262
DB 729 CGCATTTTGTGAATTTCAAGTGGCGGCAACCAATCAACTCTTCATCAATTTGCCAT 670
QY 2263 CACCGTAATGCGAGCAAAATGTCGCAATACCAATACGATAGTGCACACCGAA 2322
DB 669 CACCGTAATGCGAGCAAAATGTCGCAATACCAATACGATAGTGCACACCGAA 610
QY 2323 AGTATTAACCAATACCGAAAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATCGG 2382
DB 609 AGTATTAACCAATACCGAAAGCTGGGTTTCAAGCCCTGTGATGACAGGTTTGTATCGG 550
QY 2383 TTTGATTAATCAAAACCTTTATGCTCATATCGCAAGTCGAGAGGAGGCGAAACGAG 2442
DB 549 TTTGATTAATCAAAACCTTTATGCTCATATCGCAAGTCGAGAGGAGGCGAAACGAG 490
QY 2443 CCAAGTGTGAGCTTTCTTAATATGCAAAATCAATTTTGAAGGTAGAGAAATGACAC 2502
DB 489 CCAAGTGTGAGCTTTCTTAATATGCAAAATCAATTTTGAAGGTAGAGAAATGACAC 430
QY 2503 CGCGAAGAGATTGAAGAAATCGTGATGTGCGACTGCGGAAAGCAACGCGAACTCTTC 2562
DB 429 CGCGAAGAGATTGAAGAAATCGTGATGTGCGACTGCGGAAAGCAACGCGAACTCTTC 370
QY 2563 TTTTACGCGCATATGCGCAAGCTGACGAGATATTTGTCACAGTAATCAATATTTGGC 2622
DB 369 TTTTACGCGCATATGCGCAAGCTGACGAGATATTTGTCACAGTAATCAATATTTGGC 310
```

```
QY 2623 ACCCATCCATAGAGATGCGGGTTTCAACCGGTAAACCCAGCGCAAGACCAACAT 2682
DB 309 ACCCATCCATAGAGATGCGGGTTTCAACCGGTAAACCCAGCGCAAGACCAACAT 250
QY 2683 AATGAAATCAACCGTCTTGAAGATTGAATGAGTGGCTTGGCCACTAAACCAATCATCA 2742
DB 249 AATGAAATCAACCGTCTTGAAGATTGAATGAGTGGCTTGGCCACTAAACCAATCATCA 190
QY 2743 TCTGCAATTTGGGTGGAGCAAAATTCAAATAGAACTTTGGCTTGATGCGCGTTGCCA 2802
DB 189 TCTGCAATTTGGGTGGAGCAAAATTCAAATAGAACTTTGGCTTGATGCGCGTTGCCA 130
QY 2803 TTTAAACCGCTGCCGACCATGCGACTGCAAGAAATGTAATACAGATGAAGCCAA 2862
DB 129 TTTAAACCGCTGCCGACCATGCGACTGCAAGAAATGTAATACAGATGAAGCCAA 70
QY 2863 GTTTGCCAAGTAGGCTTTGTTGTCAGGAAATGCGGCTGAG 2909
DB 69 GTTTGCCAAGTAGGCTTTGTTGTCAGGAAATGCGGCTGAG 23
```

RESULT 2

US-10-282-122A-40936

Sequence 40936, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haebelbeck, Robert

APPLICANT: Ohlsen, Karl

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See file Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 40936

LENGTH: 852

TYPE: DNA

ORGANISM: Vibrio cholerae

US-10-282-122A-40936

Query Match

Best Local Similarity 99.3%; Pred. No. 1.2e-267;

Score 842.4; DB 13; Length 852;

```

Matches 846; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 839 GTGAAACAGTATTAGATCTTTGTGACGCCATGTCGATCAAGTGTGGTTGAAAT 898
Db 1 GTGAGACATATTATATCTTTGTGACGCCATGTCGATCAAGTGTGGTTGAAAT 60
QY 899 GAACGAACGGGCAAGCGTGTGATGCTGTGATTAATGCGGATTTGACCTACGATGTGGC 958
Db 61 GAACGAACGGGCAAGCGTGTGATGCTGTGATTAATGCGGATTTGACCTACGATGTGGC 120
QY 959 AACCAATCACTTTCTCTAGTACTACACCAAGAGTTTGGAAACCTGCGGAG 1018
Db 121 AACCAATCACTTTCTCTAGTACTACACCAAGAGTTTGGAAACCTGCGGAG 180
QY 1019 TTGCTCGGCTATATCTGTGGTACGATTAATGCGGAGTTTGGCAATTAGGTACCAA 1078
Db 181 TTGCTCGGCTATATCTGTGGTACGATTAATGCGGAGTTTGGCAATTAGGTACCAA 240
QY 1079 ACCTGGGATGCTAATGCCAATTAAACCAAGATGGCTCAACATCTTAACGTAAAGT 1138
Db 241 ACCTGGGATGCTAATGCCAATTAAACCAAGATGGCTCAACATCTTAACGTAAAGT 300
QY 1139 GAGGATGACATGGAAGCGCTGTATGCTGTTCAGGGTACGCTTGGCTAAGCTGTAGT 1198
Db 301 GAGGATGACATGGAAGCGCTGTATGCTGTTCAGGGTACGCTTGGCTAAGCTGTAGT 360
QY 1199 GGTCAATATGACAGTTGAAAAAGATTGTGATGATTGAGCGGCGGTGATGACGA 1258
Db 361 GGTCAATATGACAGTTGAAAAAGATTGTGATGATTGAGCGGCGGTGATGACGA 420
QY 1259 GGTGAATCTTAACTTACATCCGGGTGAATTTACATGGGGTGTGTCGCCCTTCG 1318
Db 421 GGTGAATCTTAACTTACATCCGGGTGAATTTACATGGGGTGTGTCGCCCTTCG 480
QY 1319 ATGTACAGCCATCTTTTCATGCTGCGGGATACCTGTATCTCAACAGTACGCGT 1378
Db 481 ATGTACAGCCATCTTTTCATGCTGCGGGATACCTGTATCTCAACAGTACGCGT 540
QY 1379 TCATGATGATGCTTGGGGTTGATTTCAACATGATGAGGTTATGTTCTTGGG 1438
Db 541 TCATGATGATGCTTGGGGTTGATTTCAACATGATGAGGTTATGTTCTTGGG 600
QY 1439 CTGATGACACATCTCAAGGAAAAAGCGGCTTGGCGTATCAAGAATCGTCAATGCG 1498
Db 601 CTGATGACACATCTCAAGGAAAAAGCGGCTTGGCGTATCAAGAATCGTCAATGCG 660
QY 1499 CACATTACCAAGATCACTCGAATTGATGCGGATGTGACGTAACGTTAGCCATT 1558
Db 661 CACATTACCAAGATCACTCGAATTGATGCGGATGTGACGTAACGTTAGCCATT 720
QY 1559 CCAGCGCTCAGTTCATATCAATCCAAAGATTAAACCTGAGAGTTTGGAACTTGG 1618
Db 721 CCAGCGCTCAGTTCATATCAATCCAAAGATTAAACCTGAGAGTTTGGAACTTGG 780
QY 1619 GTCACTTTGGATATTTTGAACGTCACCGGATATCAGTCCAGATCTTATCAATACCG 1678
Db 781 GTCACTTTGGATATTTTGAACGTCACCGGATATCAGTCCAGATCTTATCAATACCG 840
QY 1679 TTTTCAGTCTAA 1690
Db 841 TTTTCAGTCTAA 852

```

```

RESULT 3
US-10-329-960-1
; Sequence 1, Application US/10329960
; Publication No. US2003009277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PB186PI
; CURRENT APPLICATION NUMBER: US/10/329,960

```

```

; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44016)..(44016)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c

```

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119244)..(119244)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

Query Match: 25.6%; Score 744.2; DB 15; Length 1830121;
Best Local Similarity 66.7%; Pred. No. 1.4e-232;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

```
QY 2 AGAAGTTGTTATGCTCAGGGTTATCTGCAAGTTCCCAATATTGACCCCGTATTGTT 61
DB 958086 ATAGAGAAATATGAAATTAATTAATTTACTTCCCACTTGATCCAGATTTT 958145
QY 62 TCGATCGC-----CCTTAGCGGTGCGCTGTATAGCTTGATGATTTGGTGGCTTC 115
DB 958146 ACGCTTGGCGAATGAAATATGAGGCTTACCTTGATAGCTTGATGATCTTTAGCTTT 958205
QY 116 CTTTGTGATGTGTTGGCAATTCGCCAGCGATCGCGGGGAGTGTGACGCGCT 175
DB 958206 GTTTTGCACGTGGCTTCGCGTTCGCGCTAATCGCCAAATAGCGGTGACAGTA 958265
QY 176 GAGCAAGTCCTGACTTGTATTCGCGGCTTTTAGGTGATGATCGTGGCGAGTT 235
DB 958266 GATCAGTTGATGATTAATCTTTTCAACGTTTATAGGAGGCTTATTGGCGAGCTGT 958325
QY 236 GGTATGTGATCTTCAATTTGATCTGTCTTGTGACCCCTTATTATTCAAA 295
DB 958326 GGCATGATTTTCTATATCTGATCATTTCTTACAAGAACCACTTATTATTCGC 958385
QY 296 GTGTGACTGGCGGATCTCTTCACGCGGCTTATTGGGTGATACCGCATGCTC 355
```

Db 95886 GTTTGGGAAGTGAATGCTTCCAGGCTGCTTAATGCTGTAATTTGCTATGANT 958445
 Qy 356 TGGTATGCGGTGAAAAACCAACGACCTTCTTGTGTGGCCGATTTTGTGCCCCCTTA 415
 Db 958446 TGGACATCTTATCTCAAAAAGCTAATTTTGGCAACGCGCTATTTGTGGCCCTTTG 958505
 Qy 416 GTGCGATTCGCTTGGGGAATGGACGTAATGCTTAATGTAATGTAATCTTTGGGA 475
 Db 958506 ATTCGCTTGTGTTAGGTTAGCAGAAATGGTAATTTCAATCTTAATCTGAACTAATGGGA 958565
 Qy 476 CGAGTAACGGAATGCTTGGGCTTTGTAATCCCTTAATGCTGCGC---CACTGCGCGC 532
 Db 958566 CCGGAACCAATGCTTGGGCAATGATTTTCCCAATGATCCTCTTTTACGCTCGT 958625
 Qy 533 CATCTTCAAGCTTATGTAATGCTTGAAGGGGTGCTGCTTCTTATCTTAAT 592
 Db 958626 CATTCATCAACACTTATGAGCTTTTGAAGGCTGTGTGTTAAGATTCGAAT 958685
 Qy 593 TGGTATTTGTTAACTGCTGCTGCTAGGCGATGCCGATCCGATTTTACCTGGATAC 652
 Db 958686 ATTTTATTAATAAACACAGCTCAATGGCTTCTGTGCAAGTTTATCTTAATGGTTAT 958745
 Qy 653 GGTACATTCGCTTCTGTGTAATAGTCCGTGACCGAGTCTAGTTGGCTGCTT 712
 Db 958746 GGGCTTCTCGCTTTTATGGAATATGTCGTGAACCTGAAG-----TTGAATTTTC 958799
 Qy 713 GGTGCTTCAATTCATGGGGCAATCCTCTCTTCAATGCTATGCTATCTTGG 772
 Db 958800 TTTGGATTTATTAACAGAGGGCAAGCCTTGTCTTCCCAATGATTAATGCTGCTTTC 958859
 Qy 773 ATGATGTTTGTCTTCAACGCGGTTTGTATCAAGCCGTGTACAGCAAAATAGGT 832
 Db 958860 ATTTATGCTTGGCTTATTCAGC-----AAAAGTCGCTAATAAATAGGAG 958907
 Qy 833 AGTTAGGTAAACAGATTTAGATCTTTGTCAAGCGATGTCATCAAGGTGTTGGGT 892
 Db 958908 ATTTTATGAAGCAATATCTGAGCTTGTGCGCGCATTTGTAGTAAGGGGAATGGGT 958967
 Qy 893 GAAATGAAGCAACGGGCAAGCCTTGTGATCTGTATTAATGCTGATTTGACCTACAT 952
 Db 958968 GCTAATGAAGCAATGAGTAAGCAATGCTCAAGCTAATTAATGAATTAATGAT 959027
 Qy 953 GTGGGCAACATGATGTTCTCTAGTACTACACGCAAGATTTTGAAGGCTGCGTA 1012
 Db 959028 GTGGGCAATATCAATTTCCGCTGATTAATCCCGTAATTAATGGAAGGCGGAT 959087
 Qy 1013 GCGGATGCTGCTGCTAATATGCTGTGTAATCGATATGCGCGGATTTTGGCAATTAAGT 1072
 Db 959088 GCTGAATTTTATGTTATATGCTGATATGACAAATGCGCTGATTTTCCGCGCACTTGGC 959147
 Qy 1073 ACCAAAACCTGGGATGCTAATGCAATTTAAACCAAGCATGCTCAACATCCTTACCT 1132
 Db 959148 ACGAAAACCTGGGATGCTAATGCAATGGAATAAGCAAGCTTGGCTTGAATTCGATCTGT 959207
 Qy 1133 AAAGGTGAAGATGACATGGAAGCGCTGTATGCTTCAAGGTGAGCTTGGCTAAGCT 1192
 Db 959208 AAGAGGCTGATGATATGAGGCGCGCTATATGCTGTCAAGGACAGACATGCGTAAAGCT 959267
 Qy 1193 GATGTGTGCTAATTTGACCAAGTTGGAAGATTTGATGATTTGAAGCGGTGCGTGTAT 1252
 Db 959268 AATGAGAAATATATCATGATCAGTAAGTAATTTGTAATTAATTAATGAAGGATTTGAT 959327
 Qy 1253 GACCGAGGTGAATTTCTTAATCTTAATCAATCCGGGTGAATTTCAATGGGCTTTGGCG 1312
 Db 959328 GATGAGGAGAGATTTTAATCTTTTAAACCTGCGGAAATTTGATCTTGTGTCTTGT 959387
 Qy 1313 CCTTGATGTACAGCATATTTTCAATGCTGCGGATACCTTGTATCTCAACAGTACT 1372
 Db 959388 CCTTGATGTACATGATATCTTTTCTTGTGGGATACCTTGTATCTTACTACTAT 959447
 Qy 1373 CAGCGTTCAATGATGCTGCTTGGGTTGAATTTCAACATGCTGAGGTTATGTTGCTC 1432
 Db 959448 CAGCGTTCTGATGATGCTGCTGATTAATTTCAATCAATCAGGTGTTTACCTTC 959507

Qy 1433 CTTCGCTGATGGACACATCAACGGAAGGCGGCTTGGCTATCAACAGATCTGC 1492
 Db 95508 TTAGCACTTATGCTCATGATCAACGGAAGGCGGCAAGGCAATATCAATTAATTTG 959567
 Qy 1493 AATGGCACTTTACCAAGATCACTCGAATGATGCGGATGCGAGCTTAACGCTGAG 1552
 Db 95968 AATGGCACTTTATGAAGATCAGCTTGAATATGCGTGAACACTTAACGCGAG 959627
 Qy 1553 CCATTCGCGCGGCTCAGTTCATATCAATCCAAAGATTAAACACTGAGATTTGGA 1612
 Db 959628 CTTTCCATTAACCAAAATTAATCAATTAATCAATTAATCAATTAATTAATTAATTA 959687
 Qy 1613 ACTTGGGTCACTTGGATATTTTGAAGCTCACCGGATATCAGTTCACAGATCTTCA 1672
 Db 959688 ACTTGGGTCACTGATGATATTTTAAGTGTGCTATCAATCCACGACCAATTA 959747
 Qy 1673 TACCGTTTCACTAAT 1691
 Db 959748 TATCTTTTCCGCTAAT 959766

 RESULT 4
 US-10-329-670-1
 ; Sequence 1, Application US/10329670
 ; Publication No. US20040018503A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fleischmann et al.
 ; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
 ; FILE REFERENCE: PB186P1
 ; CURRENT APPLICATION NUMBER: US/10/329,670
 ; PRIOR FILING DATE: 2002-12-24
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: US 09/643,990
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/426,787
 ; PRIOR FILING DATE: 1995-04-21
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1830121
 ; TYPE: DNA
 ; ORGANISM: Haemophilus influenzae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (4747)..(4747)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (921)..(921)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (10150)..(10150)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (29298)..(29298)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (36543)..(36543)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (36551)..(36551)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (36636)..(36636)
 ; OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

```

; LOCATION: (152500) ..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530) ..(152530)

Query Match      25.6%; Score 744.2; DB 16; Length 1830121;
Best Local Similarity 66.7%; Pred. No. 1.4e-232;
Matches 1134; Conservative 0; Mismatches 538; Indels 27; Gaps 4;

QY      2 AGAAGTTGTTATGCTCAGGTTATCTGCAAGTTTCCCAATATGACCCCGATTTGTT 61
DB      958086 ATTAAGAAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 958145
QY      62 TCGATCGC-----CCTCAGCGGCGCGTGTATGCTGTATGATTAATTTGGTGGTTTC 115
DB      958146 AGCGTTGGGATAGTAATATCGCTTACGTTGATGAGCTTATGACCTTTTAAAGTTT 958205
QY      116 CTTTTGCTATGTTGTTGGCCATCGCCAGCGGATCGCGGCGAGTGTGACGCGT 175
DB      958206 GTTTTTCACGTTGCGCTTGGCGTTCGCGCTGTATATGCGCCAAATAGCGGTTGACAGTA 958265
QY      176 GACCAAGTCTGACTTGTATTCGCGCGCTTTTATGATGATCGGTGCGCGAGT 235
DB      958266 GATCAAGTTGATAGTACTTATTCACAGGTTTATGAGGAGTGTATGCGGAGCGTGT 958325
QY      236 GGTATGTGATCTTCTACAAATTTGATCTGTTCTTGTGACCTCTTATTTATTCAA 295
DB      958326 GCGGATGATTTTCTATATATCTGATCTTCTTCAAGAACCACTTATTTATTCGCG 958385
QY      296 GTGTGAGCTGCGGCGATGCTCTTCCACGCGGCGCTTATGAGTGTATCAACGCGATGTC 355
DB      958386 GTTTGGGAAGTGGAAATGCTTCCACGAGTGGCTTATGAGTATGTTGTTGATGATTT 958445
QY      356 TGGTATCGCGCTTAAACCAACGACACTTCTTGTGTGCGCATTTTGTGCCCTTAA 415
DB      958446 TGGACATCTTATCTCAAAAAGTATTTTGGCAAAAGCGCTGATTTTGTGGCGCTTG 958505
QY      416 GTGCCATTGCTTGGGGAATGGACGTAATGCTTATGATGATGTAATGTTGGGA 475
DB      958506 ATTCCTTTGTTAGTTAGGCAAGATGTAATTCATTAATCTTGAACATATGGGA 958565
QY      476 CGAGTAACGAGATGCTGCTGGCTTGTATTCCTTAATGCGGCGC---CACGCGCGCG 532
DB      958566 GCGGAAACGAATGTGCTTGGGCAATGATTTTCCGAATGATCTCTTTTACGCTGT 958625
QY      533 CATCTTCAAGCTTATGATATTCGCTTGAAGAGGCTGTTCTTATTTCTTAAT 592
DB      958626 CATCATCACACTTATTAAGCTTTTGAAGGCTGTGTGTTTACATTCGAT 958685
QY      593 TGGTTATTTGTAACCTGCTGCTAGGCAAGATTCGGAATGTTTAACTGTGATAC 652
DB      958686 ATTTTATTAATAAAACCAAGTCAATGCTTCTGTTGCAAGTTTATTTCTTAATGTTAT 958745
QY      653 GGTACATTCGCTCTCTTGTGAATACGTCCTGAGCCGAGCTCAAGTGTGCTGTT 712
DB      958746 GCGCTCTTCCGTTTATTTGGAATATGTCGTAACCTGAAC-----TTGAAAATTTTC 958799
QY      713 GGTGCTTCAATTCATGAGGCAATCTCTCTTCACTATGATGATCAATCGGATTTTG 772
DB      958800 TTTGGATTAATTAACGAGGCAAGGCTTTGCTTCCGATATATTTGGTGTGCTTTC 958859
QY      773 ATGATGTTTGTCTTACAGCGCGGTTGTATCAAGACCGGTAGACCAAAATNGGT 832
DB      958860 ATTTATGCTTGGGCTTATTCAGC-----AAAAAGCGGTCAATAAAATGAGAG 958907
QY      833 AGTTAGTGAACAGATTAATTAATCTTTGTACAGCGATGTCGATTAAGTGTGTTGGTT 892
DB      958908 ATTTTATTAAGGAATATCTTGAAGCTTTGTGCGCGGATTTGATGAAGGGAATGGGTT 958967
QY      893 GAAATGAAGCAACGCGCAAGGCTTGTGATCTGATTAATTCGATTTGACCTACGAT 952
DB      958968 GCTAATGAACGTAAGGTAAGCATTCCTCAGGCTCATTAATGACATTTAGATATGAT 959027

```

```

QY      953 GTGGCAACAATCAGTTTCTCTAGTACTACAGCAAGATTTTGGAAAAGCTCCGTA 1012
DB      959028 GTGGCAATATCAATTTCCGCTGATTAATTAATTAATTAATTAATTAATTAAT 959087
QY      1013 GCCAGTTCCTGAGCTATATTCGTGTTACGATTAATGCGCGGATTTTTCGCAATTAAGT 1072
DB      959088 GCTGATTTTATGTTATTAATTCGTGATATGATGATGATGATGATGATGATGATGATG 959147
QY      1073 ACCAAACCTGGAGATGATATGCAATTTAAACCAAGCATGCTCAACAATCTTAAGT 1132
DB      959148 ACAGAAACCTGGAGATGATATGCAATTTAAACCAAGCATGCTCAACAATCTTAAGT 959207
QY      1133 AAAAGTGAAGATGATGATGAGGAGCGGTGATGATGATGATGATGATGATGATGATG 1192
DB      959208 AAGAGCGTTGATGATATGATGAGGCGCTATATGATGATGATGATGATGATGATGATG 959267
QY      1193 GATGATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1252
DB      959268 ATGAGGAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959327
QY      1253 GACCGAGTGAATTTCTTAATCTTAACAATCGGAGTGAATTTCAATGAGGCTGTTGCGC 1312
DB      959328 GATGAGGAGATTTTAATCTTTTAAACCTGAGGGAATTTGATCTGTGCTTGTGT 959387
QY      1313 CCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
DB      959388 CTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959447
QY      1373 CAGCGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1432
DB      959448 CAGCGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 959507
QY      1433 CTTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1492
DB      959508 TTAGCATTAATGCTCAGATCAACGCGCAAAAGCGCGCAAGCATATCATTAATTTG 959567
QY      1493 AATGCGCATTTTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1552
DB      959568 AATGCGCATTTTACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 959627
QY      1553 CCATTCACGAGCGCTCAGTTCATATCAATCAATCAATCAATCAATCAATCAATCAATCA 1612
DB      959628 CTTTCCCATTAACCAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 959687
QY      1613 ACTTGGTCACTTGTGATGATTTTGAAGCTCACCGATATCAATTCACGATCTTAATCA 1672
DB      959688 ACTTGGTCAAGATGATGATTTTAAGTCTGTGCTATCAATCCACGAACCAATTA 959747
QY      1673 TACCGTTTTCAGTCTAAT 1691
DB      959748 TATCTTTTTCGTTAAT 959766

RESULT 5
US-10-282-122A-30692
; Sequence 30692, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangshu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

```

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30692
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-122A-30692
```

```

Query Match      17.4%; Score 505.4; DB 13; Length 852;
Best Local Similarity 74.6%; Pred. No. 7.8e-156;
Matches 635; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
```

```

QY      840 TGAACAGATTTAATCTTTGTCAAGCATGTCATCAAGTCTTTGGTTGAAATG 899
      2 TGAACACTATTTAACTTGGCAACGATCTGTGATGAAAGTCTTTGGTTGAAATG 61
QY      900 AACGACGGGCAAGGCTTTGACCTGATTAATGCGAATTTGACCTACATGAGGCA 959
      62 CACGACCTGGCAAGACCTTGTCTGACCGTGATTTGATGCGGATCTACCTACATGTTGCA 121
QY      960 ACAATCACTTCTCTAGTACTACACGCAAGATTTTGGAAAGCTCCGTCAGCCGAGT 1019
      122 ATTAACCAATTTCCGTTGATACACACGCAAGATTTGAAAGCGCAATTTGCAAGAT 181
QY      1020 TGCTGGCTATTTTGTGTGTTACGATATGCGCGGATTTTCCCAATTAAGTACCAAA 1079
      182 TTTTGGTTATATCCGTGGCTATGCAATGCGGCAATTTCCGCAACCTCGGCAACAAA 241
QY      1080 CTTGGGATGCTAATGCAATTTAAACCAAGCATGGCTCAACAATCTTAACGTTAAGGTG 1139
      242 CTTGGGATGCAATGCTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 301
QY      1140 AGGATGACATGGGACCGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1199
      302 CCGATGATATGGGACGCTGTTATGCGTACAAAGGCGGATGCGCAACCGCAAGCGTG 361
QY      1200 GTCATATTGACGATTTGAAAAAGATTTGATGATTTGACCGCTGCTGATGACGAG 1259
      362 AAACCGGATCAATTTGGGCAAAATGTCATTAATCTGCGCAACGCAATTTGATGACG 421
QY      1260 GGAATTTCTTAATCTTAACATCGGGGTAATTTCAATGGGGGTTTGGCGCCCTTGA 1319
      422 GGAATATCAATGCTTTCTTAACCCCGGGAATTTGAATCTGGGTGCTTACCTCTTCA 481
QY      1320 TGTACAGCATATTTTCAATGCTGAGGAGATCTTGTATCTCAACAGTACTAGCGTT 1379
      482 TGCAACCCACACCTTCTCATATCTCGGTGATACCTCTATCTACAGATTAACGCTT 541
QY      1380 CATGTGATGTGCGCTTGGGGTGAATTTCAACATGCTGCAAGTTTATGTGTTCTTGGCG 1439
      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

```

Db      542 CTGTGATGTCGCCGCTTGATTTGAATTTCAACCAATTTCAAGTGTACCTTTCTGCT 601
QY      1440 TGATGGCAACAGATCAAGGGAAGAAAGCGGCTTGCGTATACAGATCTCATATGGC 1499
      602 TGATGGCAACAAATTAACGGGCAAAAGCGGTTAAGCGGTACACAAATATCATATGAC 661
QY      1500 ACATTTACCAAGATCACTGCAATTTGATGCCGATGTGACACTTAAAGTACGATTTCC 1559
      662 ACATTTAAGAAATCACTGATCTCATGAAAAATGTGCAATTAACGTGACCCCTTCC 721
QY      1560 CAGGCGCTCACTTCATATCAATGCAATTAACAGTCAAGATTTGAAAATTGGG 1619
      722 CTTGCCAATTAATCACTTAACCCGAGATTTAACTTGAAGATTTAGAACTTGG 781
QY      1620 TCACCTTGATGATTTTGAAGCTCACCGGATATCACTTCAAGATCTTAATCAATCCGT 1679
      782 TGACATGATGATTTTAAGGTCAACGGCTATCATGACAGATGCAATTAATATCTT 841
QY      1680 TTTCACTTAA 1690
      |||||||
Db      842 TCTCTGTGTAA 852
```

RESULT 6

US-10-282-122A-22162

Sequence 22162, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu

APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlson, Kari

APPLICANT: Zyskind, Judith

APPLICANT: Wall, Daniel

APPLICANT: Trewick, John

APPLICANT: Carr, Grant

APPLICANT: Yamamoto, Robert

APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: EUTRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

PRIOR FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

SOFTWARE: PatentIn version 3.1

SEQ ID NO 22162

LENGTH: 852

TYPE: DNA

ORGANISM: Haemophilus influenzae

US-10-282-122A-22162

Query Match 15.9%; Score 462.2; DB 13; Length 852;
 Best Local Similarity 71.4%; Pred. No. 1.7e-141;
 Matches 608; Conservative 0; Mismatches 243; Indels 0; Gaps 0;

QY 840 TGAACAGATTTAGATCTTTGTCAAGCGATGTCATCAAGTGTGGTTGAAATG 899
 DB 2 TGAACCAATATCTGAGCTTTGTCCGCCATTTGTAGTAAAGGGGATGGTTGCTAAATG 61

QY 900 AACGAACGGGCAAGGCTTTGACGTGATTAAGCCGATTTGACCTACAGATGGGCA 959
 DB 62 AACGTAACGTAAGCAATGCTTCAAGCTTAAATGCAATTAAGATAGATGGCGCA 121

QY 960 ACAATCAGTTCTCTAGTACTACAGCAAGATTTTGAAGTCCGAGCGAGT 1019
 DB 122 ATAAATCAATTTCCGGGATTAATCAACCGTAAAGTTATGAAAGCGGAGTTGCGAAT 181

QY 1020 TECTGGCTATATTCGTGTATACGATATGCGCGGATTTTGGCCAAATAGTACCAAA 1079
 DB 182 TTTTAGGTTATATTCGTGTATATGACAAATGCGCTGATTTCCGCCACTTGGCAGCAAAA 241

QY 1080 CCGGAGTCTATGCAATTTTAAACCAAGCATGGCTCAACATCCTTACCGTAAAGTG 1139
 DB 242 CTGGAGATGCTATGCAATGAAATGACAGCTTGGCTTGAATCCGATCGTAGAGCG 301

QY 1140 AGGATGACATGGGACGCGTGTATGCTTCAAGGTAGAGCTTGGCTAGCGTATG 1199
 DB 302 TTGATGATATGGGCGCGTATATGCTGTCAAGGACGACATGGCTAAGCTTAATGAG 361

QY 1200 GTCATATTTGACCAATGAAAGATTTGATGATTTGAGCCGTGGCTGATGACGAG 1259
 DB 362 AAACATGATGATGACGTAACGTAATGTTAATTAAGTAAGGTAATGATGATAGAG 421

QY 1260 GTGAATTTCTTAATCTTCAATCCGGGTGATTTTCAACATGGGCTTTGCCCTTGA 1319
 DB 422 GAGAAATTTTAACTTTTAACTTGGGAAATTTGATCTTGGTGTCTTCTTGTGA 481

QY 1320 TGTACAGCATATTTTCTATGCTGGGGATACCTTGTATCAACAGTACTACAGCTT 1379
 DB 482 TGCATACGATACCTTTTCTTGTGGGATACCTTAACTTACTAGCTATCAGCTT 541

QY 1380 CATGTGATGTCCTTGGGTTGAATTTCAACATGGTGCAGTTATGCTTCTTGGC 1439
 DB 542 CCGTGTATGTCGCTTGGATTTGAATTTCAATCAATTCAGGTTTACTCTTATGAC 601

QY 1440 TGATGCGACATGTCACAGGAAAGCCGGCTTGGCTGATCAAGATGCTCATGCGC 1499
 DB 602 TTAATGCTCAGATCAAGGCAAAAGCCGCAAGGCAATATATAATTTGGAATGCGC 661

QY 1500 AATTTTACCAAGATCACTGAAATTTGATGAGGGAATGAGGCTAAAGCGTCAATTC 1559
 DB 662 ATATTTATGAAATCACTTGAATTTATGCTGACCTTAACTTAAACCGGAGCTTTCC 721

QY 1560 CAGCGCTCAGTTCAATCAATCCAAAGATTTAAACACTGCAAGATTTGAAACTTTGG 1619
 DB 722 CATTAACCAAAATTTAGAAATTAATCAAGATTAATAAGCTTGAAGTTTGAAGCTTGG 781

QY 1620 TCATCTTGGATTTTGAAGTCAAGGATATGATTCAGATCTTAAATACCGCT 1679
 DB 782 TCACATGATGATTTTAAAGTCTGCTATCAATCCACGAACCAATTAATATCTT 841

QY 1680 TTTTCACTAA 1690
 DB 842 TTTCCGTCTAA 852

RESULT 7
 US-10-282-132A-32325
 ; Sequence 32325, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangau
 ; APPLICANT: Zamudio, Cheryl
 ; APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zykied, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forayth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282.122A
 PRIOR FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 32325
 ; LENGTH: 849
 ; TYPER: DNA
 ; ORGANISM: Proteus mirabilis
 US-10-282-132A-32325

Query Match 13.9%; Score 403.8; DB 13; Length 849;
 Best Local Similarity 67.3%; Pred. No. 4e-122;
 Matches 570; Conservative 0; Mismatches 277; Indels 0; Gaps 0;

QY 840 TGAACAGATTTAGATCTTTGTCAAGCGATGTCATCAAGTGTGGTTGAAATG 899
 DB 2 TGAACCAATATCTGAGCTTTGTCCGCCATTTGTAGTAAAGGGGATGGTTGCTAAATG 61

QY 900 AACGAACGGGCAAGGCTTTGACGTGATTAAGCCGATTTGACCTACAGATGGGCA 959
 DB 62 AACGTAACGTAAGCAATGCTTCAAGCTTAAATGCAATTAAGATAGATGGCGCA 121

QY 960 ACAATCAGTTCTCTAGTACTACAGCAAGATTTTGAAGTCCGAGCGAGT 1019
 DB 122 ATAAATCAATTTCCGGGATTAATCAACCGTAAAGTTATGAAAGCGGAGTTGCGAAT 181

QY 1020 TECTGGCTATATTCGTGTATACGATATGCGCGGATTTTGGCCAAATAGTACCAAA 1079
 DB 182 TTTTAGGTTATATTCGTGTATATGACAAATGCGCTGATTTCCGCCACTTGGCAGCAAAA 241

QY 1080 CCGGAGTCTATGCAATTTTAAACCAAGCATGGCTCAACATCCTTACCGTAAAGTG 1139
 DB 242 CTGGAGATGCTATGCAATGAAATGACAGCTTGGCTTGAATCCGATCGTAGAGCG 301

QY 1140 AGGATGACATGGGACGCGTGTATGCTTCAAGGTAGAGCTTGGCTAGCGTATG 1199
 DB 302 TTGATGATATGGGCGCGTATATGCTGTCAAGGACGACATGGCTAAGCTTAATGAG 361

QY 1200 GTCATATTTGACCAATGAAAGATTTGATGATTTGAGCCGTGGCTGATGACGAG 1259
 DB 362 AAACATGATGATGACGTAACGTAATGTTAATTAAGTAAGGTAATGATGATAGAG 421

QY 1260 GTGAATCTTAACTTCTACATCCGGGTGAATTTACATGGGGTGTGGCCCTTCA 1319
 DB 422 GTGAGATAGTCACTTTTATATATCCGGAGAAACAGCGTTAGGGTGTAGTCATCA 481
 QY 1320 TGTACAGCCATCTTTTTCATTGCTGGGGGATACCTTGTATCTCAACAGTACAGCTT 1379
 DB 482 TGCATACACATACCTTTTTCATTGCTGGGTGACAGGCTTATTTAACTCATATCAACGTA 541
 QY 1380 CATGATGATGCGCTTGGGGTGAATTTCAACATGATGAGGTTATGTTGTTCCCTGGGC 1439
 DB 542 GCTGTGATGCTCCATTTAGGGTGTGAATTTTAAATTCATAGCTTGTATTTATTAAGCGC 601
 QY 1440 TGATGACACAGATCAAGGAGAAACCGGGCTTGGCGATCAAGATCGTCAATGCGC 1499
 DB 602 TGTGGCTCAATACAGGGGATTAACAGGTAAAGCATTCATAGATGATTAATGCTC 661
 QY 1500 ACATTACCAAGATCACTCGAATGATGCGGATGTCAGCTAAACGTGAGCCATTCC 1559
 DB 662 ACATTATGAGAACCAATTAACCTGATGCGGATGTCATTAATAAGAGCCATTGC 721
 QY 1560 CAGGCGCTCAGTTCATTCATCCAAAGATTAAACACTGAGGATTTGGAACTTGGG 1619
 DB 722 CATTAAGCTACATATTAATCCAAACATTTAAGACATTTAGATGATTAAGAACTTGGG 781
 QY 1620 TCACCTTGTGATTTTGTACGCTCACCGGATACAGTTCAGGATTCATTCATACCCGT 1679
 DB 782 TCAGACAGATGACCTTGTAGTTGAAGTTTACATGTCACAGAGCGATTAATAATCCGT 841
 QY 1680 TTTTCACT 1686
 DB 842 TCACCGT 848

RESULT 8

US-09-741-669-182
 ; Sequence 182, Application US/09741669
 ; Patent No. US2002022718A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Forsyth, R. Allyn
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; TITLE OF INVENTION: Genes identified as required for
 ; proliferation of *E. coli*
 ; FILE REFERENCE: ELITRA.009A
 ; CURRENT APPLICATION NUMBER: US/09/741,669
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: US 60/173005
 ; PRIOR FILING DATE: 1999-12-23
 ; NUMBER OF SEQ ID NOS: 481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 182
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: *Escherichia coli*
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(876)
 US-09-741-669-182

Query Match 8.5%; Score 248.2; DB 9; Length 876;
 Best Local Similarity 58.8%; Pred. No. 1.7e-70;
 Matches 510; Conservative 0; Mismatches 288; Indels 69; Gaps 2;
 QY 14 ATGCTCAGGGTATCTGCAAGTTCCCAATATGACCCCGATGTTTTCATGCGCCCT 73
 DB 1 ATGACACAGTATCTCATTTCCGAGATTGATCCGGTATCTCAATAGGACCC 60
 QY 74 CTAGCGGTGCGGTGATGCTGATGATTTGGGGTTCCTTTTGCATAGGCTG 133
 DB 61 GTGGGCTTCACTGTTAGCGGCTGATGATCTGGGGTTCATTTTGCATAGGCTG 120
 QY 134 GCCAATCCGAGCGGATCGCGGGCAGTGTGGAAGCGTGAAGACTCTGACTTG 193

DB 121 GCAACACGACCGGGGAATCGCCGGGACGCGCTGACCAAAAATGAAGTGAATAACTTA 180
 QY 194 TTAATCGCGGCTTTTAAAGTATGATGATCGGTGGCCAGTGTGATGATCTTAC 253
 DB 181 CTATAGGGGCTTCTCGGCGTCTTCTCGGGGACATATGGTATGTTCTGTTTAC 240
 QY 254 AATTTGATCTTCTTGTGACCCCTTTATTTATTCAAAGTGGAGTGGCGGATG 313
 DB 241 AATTTCCGAGTTTATAGCCGATCCGCTGATCTGTTCCGTTCTGGACGCGGATG 300
 QY 314 TCTTTCACGCGCGCTTATGGGTGATACCGCATGTTCTGGTATGCGGTAAC 373
 DB 301 TCTTTCACGCGCGCTGATGGGTTATCGTGTGATGATTAATTCGCGCGGCTACT 360
 QY 374 CAAGCACCTCTTTGATGCGGATTTGTTGCCCCCTTATGACCATGCTTGGGG 433
 DB 361 AAAGTCTTCTTTCAGAGTCTGATTTTATGACACATCATATCCGTTGGTGT 420
 QY 434 ATGGACGTATCGGTAACTTTATGAATAGTAACTTTGGGACGAGT----- 480
 DB 421 GCCGGGCTTGGGCACTTATTAACGTTAAATGTTGGGCGCGGTTGACCCGAATTC 480
 QY 481 -----AAGGAT 487
 DB 481 CCGTTTGCATGCTGTTCCCTGCGCTCCGTAACAGAAATATTTGCTGCAACAC 540
 QY 488 GTGCGTGGGCTTTTATTCCTATAGTGCCGACATCGCGGCAATCTTCAAGCTT 547
 DB 541 CCGAGTGGCAATCATTTTGACACATTAAGCTGTGCTGCGCGCACCACTCAAGCTT 600
 QY 548 TATGAATTCGCTTAGAAGCGGTGCTGTTCTTTATTTCTTAATGGTATTTGTA 607
 DB 601 TACGAGCTGCTGGAAGGTGTGTGCTTTATTAATCTCAACCTGATATATCGTAA 660
 QY 608 CCTGTCCGCTAGGACGAGTATCCGAGCTGTTTATGCTGATACGTTATCCGCTTC 667
 DB 661 CCAGCCCAATGGGAGCTGCTCAGGTTGTTCTGATGTTGTTACGCGCGCTTCGACATC 720
 QY 668 CTGTGGAATACGTCGCGGACCGAGATGCTAGTGTG---GGTCTGTTGGTGGCTCAT 724
 DB 721 ATGTGATGTTTTCGCGACGCGGACGATTTACCGGTGCTGGGTGACGATATC 780
 QY 725 TCAATGGGCAAAATCTCTCTTACTATGATGATCAATCGGTATTTGATGATGATG 784
 DB 781 AGCATGGGCAATTTCTTTCATCCGATGATTTGCGGGTGTGATCATATGATGCTGG 840
 QY 785 TCTTACAGCGCGGTTGTATCAAGAC 811
 DB 841 GCATATGCTCGCACCCACAGCAACAC 867

RESULT 9

US-09-790-988-1
 ; Sequence 1, Application US/09790988
 ; Patent No. US20020127687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHIGENOBU, SHUJI
 ; APPLICANT: MATSUMAE, HIDEKI
 ; APPLICANT: HATTORI, YOSHIO
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
 ; FILE REFERENCE: 081356/0159
 ; CURRENT APPLICATION NUMBER: US/09/790,988
 ; CURRENT FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: JP2000-107160
 ; PRIOR FILING DATE: 2000-04-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 640681
 ; TYPE: DNA
 ; ORGANISM: *Buchnera* sp.

US-09-790-988-1

Query Match 5.2%; Score 152.2; DB 9; Length 640681;
 Best Local Similarity 52.6%; Pred. No. 1,7e-36;
 Matches 432; Conservative 0; Mismatches 323; Indels 66; Gaps 2;

```

QY 35 TTTCCTAATATGACCCCGCTATGTTTGGATCGCCCTTACGCGGTGCGGTGATGCG 94
DB 479304 TTTCCTAATATGACCCCGCTATGTTTGGATCGCCCTTACGCGGTGCGGTGATGCG 94
QY 95 TTGATGATGTTGGTGGTTCCTTTTGGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 154
DB 479364 TTGATGATGTTGGTGGTTCCTTTTGGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGG 154
QY 155 GCGGCGAGTGTGACGCGGTGACGAGTGTGACGAGTGTGACGAGTGTGACGAGTGTGACG 214
DB 479424 AATTAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 214
QY 215 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 274
DB 479484 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 274
QY 275 GACCTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 334
DB 479544 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 334
QY 335 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 394
DB 479604 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 394
QY 395 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 454
DB 479664 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 454
QY 455 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
DB 479724 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
QY 512 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 517
DB 479784 AATTTCTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479843
QY 518 -----GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
DB 479844 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
QY 569 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
DB 479904 ATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
QY 629 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
DB 479964 AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688
QY 689 CCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
DB 480024 CCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748
QY 749 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789
DB 480084 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 789

```

RESULT 10
 US-10-672-787-35
 ; Sequence 35, Application US/10672787
 ; Publication No. US20040067554A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LAGACE, Robert, E.
 ; APPLICANT: PATTERSON, Chandra
 ; APPLICANT: BERG, Kim, L.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
 ; FILE REFERENCE: ELITRA.025C1

;; CURRENT APPLICATION NUMBER: US/10/672,787
 ;; CURRENT FILING DATE: 2003-09-26
 ;; PRIOR APPLICATION NUMBER: 09/596,002
 ;; PRIOR FILING DATE: 2000-06-16
 ;; NUMBER OF SEQ ID NOS: 41
 ;; SOFTWARE: PERL Program
 ;; SEQ ID NO 35
 ;; LENGTH: 96109
 ;; TYPE: DNA
 ;; ORGANISM: Moraxella catarrhalis
 US-10-672-787-35

Query Match 4.9%; Score 142.2; DB 13; Length 96109;
 Best Local Similarity 59.1%; Pred. No. 8.5e-34;
 Matches 262; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

```

QY 38 CCATATTTGACCCCGTATGTTTGGATGATGATGATGATGATGATGATGATGATGATGATG 97
DB 51130 CCATATTTGACCCCGTATGTTTGGATGATGATGATGATGATGATGATGATGATGATGATG 97
QY 98 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 157
DB 51190 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 157
QY 158 GCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 217
DB 51247 GCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 217
QY 218 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277
DB 51307 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 277
QY 278 CCTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 337
DB 51367 CCTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 337
QY 338 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 397
DB 51427 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 397
QY 458 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
DB 51547 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

```

RESULT 11
 US-10-297-465A-1/c
 ; Sequence 1, Application US/10297465A
 ; Publication No. US20040142413A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Simpson, Andrew
 ; APPLICANT: Reinach, Fernando
 ; APPLICANT: Setubal, Joao
 ; APPLICANT: Medeiros, Paulo
 ; TITLE OF INVENTION: Isolated genome of *Xylella fastidiosa* and uses thereof
 ; FILE REFERENCE: PAPSP.202 US (10213376)
 ; CURRENT APPLICATION NUMBER: US/10/297,465A
 ; CURRENT FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: PCT/IB01/01618
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: 60/209,906
 ; PRIOR FILING DATE: 2001-06-17
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 2731748
 ; TYPE: DNA
 ; ORGANISM: *Xylella fastidiosa*

; PRIOR FILING DATE: 2001-11-01
 ; PRIOR APPLICATION NUMBER: US 60/330,462
 ; PRIOR FILING DATE: 2001-10-22
 ; Remaining Prior Application data removed - See file wrapper or PAM.
 ; NUMBER OF SEQ ID NOS: 2221
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1988
 ; LENGTH: 3950
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. NM_053380
 US-10-152-319A-1988

Query Match 2.7%; Score 77.8; DB 12; Length 3950;
 Best Local Similarity 56.7%; Pred. No. 1.9e-13;
 Matches 164; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

QY 2480 TCTAGAGTAGAAGATCAGACCGCGAGAAGATTGAAAAATCGTGATGTGGCACTG 2539
 DB 724 TCCAGAGGTAAAGACCAACAGAGAGCCAGTTGAAGAAGTCAGACAGTCTCT 665
 QY 2540 GCGAAGACGCGCGAAATCTTCTTTACAGCCGATATGSCCAAGCTGACAGAGATTG 2599
 DB 664 GCAAAATGCCCTTCTGAACCTCATTTCTCTCTGCTGCATGAATGSCCAAAATCGTTG 605
 QY 2600 GTCACAGTAGTACCATATTTGGACCCATCATCATAGTAAATGCGGTTTCAAC---CGGT 2656
 DB 604 GTGATGAGAGTCCCGATGTTAGACCCATGATGATGGATGGACGCGGACTGACAGC 545
 QY 2657 AACCCACGCGAAGACAGACCAACATATATAGAAAGTACCGCTTGAGATTGAATCACT 2716
 DB 544 AAGAGAGAGCAACATGCTGAGAGATGAGAGAGAGTGTGAGCTTGACACCATG 485
 QY 2717 GCCGTTCACCTAAACCAATCATCATCTCTGCAATTGGGTGGAGCA 2765
 DB 484 ACTGTCACGACGACCCCTATCAAGTCCAGCCACGAGATTGACATTA 436

RESULT 14
 US-10-276-774-1331/c
 ; Sequence 1331, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tan, Y. Tom et al
 ; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21272-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1331
 ; LENGTH: 2208
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-276-774-1331

Query Match 2.6%; Score 75.4; DB 13; Length 2208;
 Best Local Similarity 56.0%; Pred. No. 7.6e-13;
 Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTAGAAGATCAGACCGCGAGAAGATTGAAAAATCGTGATGTGGC 2535
 DB 738 CACCTTCACGCGGAGAGACCAACAGAGCCAGTTGAAGAATCATGACAGTGGC 679
 QY 2536 ACTGCGAAGACGCGGAAATCTTCTTTACAGCGCATATGSCCAAGCTGACAGAGT 2595
 DB 678 TCTGCAAAAGCTTCTTCACTCACTCGATCTCCCACTGATGAGCGCAACATAGT 619

QY 2596 ATTGGTCACAGTAGTACCAATATTGGACCCATCACATAGGAATGCGGTTTCAACCG- 2654
 DB 618 GTTGGATGTAGACGTTCCATTTGGCCCCCATGATTAATGGAGATGGACCCGAAACGT 559
 QY 2655 --GTAACCCACGCGAAGACAGCAACAAATATAGAAAGTACCGCTTGAGATTGAAT 2712
 DB 558 GAGCAATAGAAAGAGACACCATGCTGCAACAGTAGACGTTGAGTGTGAGACTCTGCAC 499
 QY 2713 CAGTCCGCTTGACCTAAACCAATCATCATCTTGCAATTGGGTGGAGCA 2765
 DB 498 CAAGACGCTACACACACCCGATCATCAAGCCCAACAAAGGTTGACATTA 446

RESULT 15
 US-10-173-999-149/c
 ; Sequence 149, Application US/10173999
 ; Publication No. US20040005563A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
 ; TITLE OF INVENTION: And Methods of Screening for Modulators of Ovarian
 ; FILE REFERENCE: 018501-002420US
 ; CURRENT APPLICATION NUMBER: US/10/173,999
 ; CURRENT FILING DATE: 2002-06-17
 ; PRIOR APPLICATION NUMBER: US 60/299,234
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: US 60/315,287
 ; PRIOR FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/372,246
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 163
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 149
 ; LENGTH: 2280
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-173-999-149

Query Match 2.6%; Score 75.4; DB 16; Length 2280;
 Best Local Similarity 56.0%; Pred. No. 7.7e-13;
 Matches 164; Conservative 0; Mismatches 126; Indels 3; Gaps 1;

QY 2476 CATTTCTAGAGTAGAAGATCAGACCGCGAGAAGATTGAAAAATCGTGATGTGGC 2535
 DB 771 CACCTTCACGCGGAGAGACCAACAGAGCCAGTTGAAGAATCATGACAGTGGC 712
 QY 2536 ACTGCGAAGACGCGGAAATCTTCTTTACAGCGCATATGSCCAAGCTGACAGAGT 2595
 DB 711 TCTGCAAAAGCTTCTTGAACCTCATCTCGATCTCCCACTGATGAGCGCAACATAGT 652
 QY 2596 ATTGGTCACAGTAGTACCAATATTGGACCCATCACATAGGAATCGCGCTTCAACCG- 2654
 DB 651 GTTGGTGAATGACGTTCCAAATGTTGGCCCCCATGATTAATGGAGAGCGCCGAACGT 592
 QY 2655 --GTAACCCACGCGAAGACAGCAACAAATATAGAAAGTACCGCTTGAGATTGAAT 2712
 DB 591 GAGCAATAGAAAGAGACACCATGCTGCAACAGTAGACGTTGAGTGTGAGACTCTGCAC 532
 QY 2713 CAGTCCGCTTGACCTAAACCAATCATCATCTTGCAATTGGGTGGAGCA 2765
 DB 531 CAAGACGCTACACACCCGATCATCAAGCCCAACAAAGGTTGACATTA 479

Search completed: July 31, 2004, 19:10:09
 Job time : 903 secs

This Page Blank (Unintentional)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 12:17:10 ; Search time 4825 Seconds

(without alignments)
18003.955 Million cell updates/sec

Title: US-09-700-712a-1

Perfect score: 2909
Sequence: 1 gagagaagcttctatgcctc.....cagcgaatcgcgctgcag 2909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_escba:*
2: em_escba:*
3: em_escba:*
4: em_escba:*
5: em_escba:*
6: em_escba:*
7: em_escba:*
8: em_escba:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_esc3:*
12: gb_esc4:*
13: gb_esc5:*
14: gb_esc6:*
15: em_escba:*
16: em_escba:*
17: em_escba:*
18: em_escba:*
19: em_escba:*
20: em_escba:*
21: em_escba:*
22: em_escba:*
23: em_escba:*
24: em_escba:*
25: em_escba:*
26: em_escba:*
27: em_escba:*
28: gb_esc1:*
29: gb_esc2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249.8	8.6	1128	28	BZ577690 msh2_5528
2	194.8	6.7	805	28	BZ556310 pac81-60
3	191.8	6.6	853	28	BZ578932 msh2_6070
4	161	5.5	1141	28	BZ559267 pac82-164

Result No.	Score	Query Match	Length	ID	Description
5	152	5.2	1375	28	BZ553447 pac81-60
6	115.6	4.0	406	29	U83227 Pseu
7	111.2	3.8	1034	28	BZ561697 pac82-164
8	111	3.8	878	14	CD379073 PTM004242
9	111	3.8	921	14	CD382942 PTM008112
10	106.2	3.7	855	14	CD376447 PTM01614
11	104	3.6	843	14	CAS88286 hab49d11
12	103.8	3.6	813	14	CAS88344 hab50c01
13	103.2	3.5	828	14	CA475083 AGNCOURT
14	96	3.3	685	13	BM039803
15	96	3.3	686	13	BM039194
16	96	3.3	801	14	CAS88156 hab47903
17	92.4	3.2	662	14	CAS88861 hab39c07
18	89	3.1	522	13	BM061136
19	89	3.1	632	13	BM295886
20	88.8	3.1	524	14	CD336403
21	86.6	3.0	591	12	BG732886
22	85	2.9	683	10	AM423104
23	84.8	2.9	700	14	CA531127
24	84.8	2.9	712	14	CA373071
25	84.6	2.9	483	29	CG637075
26	84.6	2.9	606	10	AM546106
27	84.6	2.9	668	13	BY746793
28	84.6	2.9	976	13	BY746114
29	84.6	2.9	2255	11	AK088606
30	84.6	2.9	2256	11	AK088134
31	84.6	2.9	3128	11	AK03663
32	84	2.9	677	14	CB420463
33	83.8	2.9	756	12	BI078307
34	82.8	2.8	776	14	CF995901
35	82.4	2.8	528	14	CD330233
36	82	2.8	748	13	BU055160
37	81.8	2.8	673	14	CD313851
38	81.4	2.8	407	14	CD330290
39	81.4	2.8	1121	11	AK011435
40	80.8	2.8	945	11	BF783860
41	80.8	2.8	2486	11	AK004832
42	80.2	2.8	952	10	BF782519
43	80.2	2.8	1024	12	BM924499
44	79.8	2.7	974	10	BF607905
45	79.2	2.7	567	9	AM108416

ALIGNMENTS

RESULT 1
BZ577690
LOCUS
DEFINITION msh2_5528.y2 msh Pseudomonas aeruginosa genomic clone msh2_5528,
genomic survey sequence.

ACCESSION BZ577690
VERSION BZ577690.1 GI:27212751

KEYWORDS
SOURCE GSS.

ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

REFERENCE 1 (bases 1 to 1128)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press

JOURNAL COMMENT
Contact: Chris K. Raymond
Genome Center

University of Washington
Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954
Fax: 2066857244

Email: ckraymond@u.washington.edu
Classes: shotgun.

FEATURES
Location/Qualifiers


```

source
1..1128
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="mh2_5528"
/clone_1ib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."
ORIGIN
Query Match      8.6%; Score 249.8; DB 28; Length 1128;
Best Local Similarity 59.7%; Pred. No.1,9e-61;
Matches 453; Conservative 0; Mismatches 301; Indels 5; Gaps 2
OY      29 CTGCAGTTTCCCAATATTGACCCCGTATTTGTTGATCGGCGCTCTAGCGGTGGCTGG 88
Db      93 CTGACGTATCCCAATCGATCCGGTTCGCTGGCCATCGGCCCTGAAGATCACTGG 152
OY      89 TATGGCTGATGATTTGGTGGGTTTCCCTTTTGTATGTTGGTTGGCAATGGCGAGG 148
Db      153 TACGGCTGATGATCTTAATCGGAATCGGCGCGCTGGCTGTGGCTGGCGCGGATG 212
OY      149 GATCGCGCGGCGAGTGTGACGCGTGAAGCAAGTCTGACTTTATTCGCGCGCTTT 208
Db      213 AAGCGCTTGCACCCGACCTGCAGCAAGAGCGCCCTTTCGACTGTGTCTTCTGGGTCCGC 272
OY      209 TTAGGTGATGATCGGTGGCCGAGTTGGTATGATCTTCTCAATTTTGATCTGTTTC 268
Db      273 TGCGGCGTATCTCGGTGGCGCGCTGGGCTACGTGTTCTCAACCTGGACGATAC 332
OY      269 CTGTGTGACCTCTTATTTATTCAAATGTTGGAATGCGCGGANTCTCTTCACGCGCGC 328
Db      333 ATCGCAACCCGACGCTGATCTTCAGAGTCTGGAAGCGGATGATCTTCTCATGGCGCG 392
OY      329 TTATTGGGTGTGATCACCGGCATGTTCTGTATGCGCGTAAACCAACGCACTTCTTT 388
Db      393 CTGCTCGGCGTGTACTGTGGCGGTCTGTGTGTTCCGGCAAGCGCATGCGAAGACTTCTTC 452
OY      389 GGTGTGGCCGATTTTGTGGTCCCTTTAGTCCATTGGTTTGGGANTGGAGTATCGGT 448
Db      453 CAGCTGATGACTTATCATCGCCCGCTGGTGGCCATCGGCTGGCGCGGCGCATCGAGC 512
OY      449 AACTTATGATATGGAACCTTTGGGAGAGATACAGATGTGCCCTTGGGGCTTTTGTATTC 508
Db      513 AACTTCATCACTCGGAACGTGGGGCAAGGTCAACGATGTGGCCCTGGGCGCAATGTCTTC 572
OY      509 CCTAATGTGGGCGCACTGCGCGGCATCCTTCACAGCTTTATGAAATTCGCTTAGAAGAC 568
Db      573 CCCAAGCGGCGCGCGCTGCGCGGATCCTTCGACGCTGTACCAATTTGACCTTGAAGGC 632
OY      569 GTGTGTTGTCTTATTTCTTAATTTGTTATTTGGTAAACCTGTCGCGCTAGCAGCGTA 628
Db      633 GTGCGCGTGTGTCATCTCTGCGCTGGTGCACCCGCAAGACACGCGCGCATCGGTC 692
OY      629 TCCGGACTGTTTATGCTGTGATACGCTACATTCGCGCTTCTTGTGTGAATAGTCCGTGAG 688
Db      693 TC--GGCTGTGTCGTGTGCTCGCGGATCTTCGCTTGTGCTGTGAATTCGTGCGGTG 750
OY      689 CCAGATGCTCAGTTGGG--TCGTGTTGGTGGCTCATTTCAATGGGGCAATTCCTCTCC 745
Db      751 CCGATGCGCACTCGGCTACTCTGCGCTGGGGTTTGTCTGACCATGGGCGCAAGTGTCTTCG 810
OY      746 TTACTATGATGATCATCGGATTTTGTGATGATGTTTGG 784
Db      811 GTACCGATGTGTCTGGCGGCAATTGGCTGATGTCTGG 849

```

VERSION	B2556310	GI:27166959
KEYWORDS	Pseudomonas aeruginosa	
SOURCE	GSS.	
ORGANISM	Pseudomonas aeruginosa Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.	
REFERENCE	1 (bases 1 to 805) Spencer, D. H., Raymond, C. K., Smith, E. E., Sims, E. E., Hastings, M., Burns, J. L., Kaul, R. and Olsen, M. V. Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library J. Bacteriol. (2002) In press	
JOURNAL	Contact: Chris K. Raymond	
COMMENT	Genome Center University of Washington Box 352145, Seattle, WA 98105-2145, USA Tel: 2062216954 Fax: 2066857244 Email: ckraymond@u.washington.edu Class: shotgun.	
FEATURES	source Location/Qualifiers 1..805 /organism="Pseudomonas aeruginosa" /mol_type="Genomic DNA" /strain="1-60" /db_xref="taxon:287" /clone="pacsl-60_5617" /clone_11b="pacsl-60" /note="clinical isolate 1-60 Whole genomic shotgun library."	
ORIGIN	Query Match 6.7%; Score 194.8; DB 28; Length 805; Best Local Similarity 60.6%; Pred. No. 2.5e-45; Matches 319; Conservative 0; Mismatches 207; Indels 0; Gaps 0;	
QY	47	GACCCCGATTGTTTGCATGCGCCCTACGCGGCGCTGATAGCTGATTTG 106
DB	103	GATCTGTGGCTGCGCCATCGGCCGCTGAAGATCACTGACGGGCTATGACTG 162
QY	107	GTGGGTTTCTTTTGTCTATGTGTGGCCATCGCCGAGCGATCGCGCGGCAATG 166
DB	163	ATCGGAGATCGCGCGGCGCTGCTGTGCGCTGCGCGGATGAAAGCGTTGACCCGACC 222
QY	167	TGACGCGTGAGCAAGTCTGACTTGTATTTCGCGGCTTTTAAAGTGAATGATCG 226
DB	223	TGAAACAAAGAGCGGCTTTCGACCTGGTCTTTCGGGTCGCGCTGCGGCTGATCCTCG 282
QY	227	GCGCGAGTTGATATGATCTTTCACAAATTTATCTGTTCTTGTGACCCCTTTAT 286
DB	283	GCGCGCTGCGCTACGTCGTGTTTACAACTCGACGAGTACATCGCAACCCGACGCTG 342
QY	287	TTAATCAAGTGTGATCGCGGCGCATGCTTCCACGCGGCTTATTTGGTGTGATCAC 346
DB	343	ATCTTCAGAGTCTGGAAGAAGCGGATGTCCTTCATGCGCGGCTGCTGGGCTGATGCTG 402
QY	347	GCGATGTTCTGATGCGCGTAAAAACCAACGACCTTCTTGTGTCGCGGATTTGTT 406
DB	403	GCGGCTGTGTGTGCGCAAGCGCATGCGCAAGAGCTTCTTCAGACTGATGACTTCATC 462
QY	407	GCGCTTTAGTGCCATTGGGTTTGGGGATGGGACGTATCGGTAACTTATGATATGTAA 466
DB	463	GCGCGCTGTCGCGCATCGGCTGGGCGCGCGGCGCATCGGCACTTATCAACTCGAAA 522
QY	467	CTTTGGGGAGCAGTAAAGGATGTGCTTGGGCTTTTGTATCCCTTAATGATGTGCGCCACTG 526
DB	523	CTGTGGGCAAGTACGCGATGTGCTTGGGCTCATGTCCTTCCCAAGCGGCGCCGCTG 582
QY	527	CGCGGCATCCTTACAGCTTATCAATTCGCTTTAGAAAGCGTGG 572
DB	583	CGCGCATCCTTCCGACCTGTACAGTTCGCGCTGAAAGCGTGG 628

```

RESULT 3
BZ578932      853 bp   DNA      linear   GSS 17-DEC-2002
LOCUS         msh2_6070.y2 msh Pseudomonas aeruginosa genomic clone msh2_6070,
DEFINITION    genomic survey sequence.
ACCESSION     BZ578932
VERSION       BZ578932.1   GI:27213993
KEYWORDS      GSS.
SOURCE        Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE     1 (bases 1 to 853)
AUTHORS      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Haastings,M.,
              Burns,J.L., Kaul,R. and Olsen,M.V.
              Whole-genome-sequence variation among multiple isolates of
              Pseudomonas aeruginosa library
              J. Bacteriol. (2002) in press
              Contact: Chris K. Raymond
              Genome Center
              University of Washington
              Box 352145, Seattle, WA 98105-2145, USA
              Tel: 2062216954
              Fax: 2066857244
              Email: craymond@u.washington.edu
              Class: shotgun.

FEATURES
  source
    1..853
    /organism="Pseudomonas aeruginosa"
    /mol_type="genomic DNA"
    /strain="MSH"
    /db_xref="taxon:287"
    /clone_1fb="msh2_6070"
    /clone_1lb="msh"
    /note="Environmental isolate. Whole genomic shotgun
    library."

ORIGIN
  Query Match      6.6%; Score 191.8; DB 28; Length 853;
  Best Local Similarity 60.0%; Pred. No. 2e-44;
  Matches 336; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY 29 CTGCACTTCCCAATATGACCCCGATTTGTTTGCATGCGCCCTAGCGGCGCTG 88
DB 168 CTGACGTATCCCAATGATCGATCGGTTGCGCTGCGCCATCGCCCGCTGAATCACTG 227
QY 89 TATGCTTGATGATTTGGTGGGTTCTTTTGTATGATGATGATGATGATGATGATGATG 148
DB 228 TAGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 287
QY 149 GATCGCGCGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 208
DB 288 AAGCGCTTGACCGGACCTGACCAAGAGCGCTTCCGACCTGCTTCTGCGTGGCC 347
QY 209 TTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
DB 348 TCGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 407
QY 269 CTGTGACCTCTTATTTATTTCAAGTGTGATGATGATGATGATGATGATGATGATG 328
DB 408 ATCGCAACCGGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 467
QY 329 TTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 388
DB 468 CTGCTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 527
QY 389 GGTGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
DB 528 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 587
QY 449 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 508

```

```

DB 588 AACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647
QY 509 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 568
DB 648 CCCAAGCGGCGGCTGCGGCGGATGATGATGATGATGATGATGATGATGATGATG 706
QY 569 GTGCTGCTGCTTATTTCT 588
DB 707 CGTGGCGCTGCTGATCTCT 726

RESULT 4
BZ559267      1141 bp   DNA      linear   GSS 17-DEC-2002
LOCUS         pacs2-164_1355.s1 pacs2-164 Pseudomonas aeruginosa genomic clone
DEFINITION    pacs2-164_1355, genomic survey sequence.
ACCESSION     BZ559267
VERSION       BZ559267.1   GI:27175318
KEYWORDS      GSS.
SOURCE        Pseudomonas aeruginosa
ORGANISM      Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE     1 (bases 1 to 1141)
AUTHORS      Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Haastings,M.,
              Burns,J.L., Kaul,R. and Olsen,M.V.
              Whole-genome-sequence variation among multiple isolates of
              Pseudomonas aeruginosa library
              J. Bacteriol. (2002) in press
              Contact: Chris K. Raymond
              Genome Center
              University of Washington
              Box 352145, Seattle, WA 98105-2145, USA
              Tel: 2062216954
              Fax: 2066857244
              Email: craymond@u.washington.edu
              Class: shotgun.

FEATURES
  source
    1..1141
    /organism="Pseudomonas aeruginosa"
    /mol_type="genomic DNA"
    /strain="2-164"
    /db_xref="taxon:287"
    /clone_1fb="pacs2-164_1355"
    /clone_1lb="pacs2-164"
    /note="clinical isolate 2-164 Whole genomic shotgun
    library."

ORIGIN
  Query Match      5.5%; Score 161; DB 28; Length 1141;
  Best Local Similarity 64.6%; Pred. No. 2.7e-35;
  Matches 272; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY 380 ACCCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 439
DB 628 AGCTTCTTCCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 569
QY 440 GGTATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 499
DB 568 CGCATCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 509
QY 500 TTTGATTCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 559
DB 508 ATGCTTCTTCCCAAGCGGCGAC-GTGCAGCGCATCTCTGACGCTGATCAAGTTG 450
QY 560 TTAGAAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 619
DB 449 CTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 390
QY 620 GCGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 679
DB 389 GCTGCTGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330

```

```

Qy 680 GTCCGTGAGCCAGATGCTCAGTTGGG--TCTGTTGGTGCTTCATTTCAATGGGCGAA 736
Db 329 GTCCCGCGTCCGGATGACCAGCTCGGCTGAGCTGCTGGGCTTGGCTGACCATGGGCGAG 270
Qy 737 ATCTCTCTCTTACCTATGATGATCATCGGTATTTTGATGATGATGTTTGTCTTACAGGCG 796
Db 269 GTGCTCTGTGATCCGATGCTGCTGGCGGATTCCTGATGCTGCTGAGCGCGCGC 210
Qy 797 G 797
Db 209 G 209

RESULT 5
BZ53447 1375 bp DNA linear GSS 17-DEC-2002
LOCUS pacel-60_4149, genomic survey sequence.
DEFINITION BZ53447
VERSION BZ53447
KEYWORDS BZ53447.1 GI:27159498
SOURCE GSS.
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1375)
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Pseudomonas aeruginosa library
J. Bacteriol. (2002) in press
CONTACT: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source location/Qualifiers
1..1375
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacel-60_4149"
/clone_lib="pacel-60"
/note="Clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Query Match 5.2%; Score 152; DB 28; Length 1375;
Best Local Similarity 57.7%; Pred. No. 1.3e-32;
Matches 291; Conservative 0; Mismatches 210; Indels 3; Gaps 1;
Qy 270 TTGCTGACCCCTTATTATTATCAAGTGTGAGCTGCGGCGCATGCTTCCACGGCGGT 329
Db 84 TTCCTGACGCCCCGCGATTTCTAGTCTGCAAGGGCGGCGATGCTTCCATGCGGCG 143
Qy 330 TATTGGGTGTATCACCGGCATGTTCTGATGCGCGTAAACCAAGCGACCTTCTTTG 389
Db 144 TGCTGGGGCTGATGCTGCGGTCTGATGCTTGGCAAGCGCCATGCGAAGCTTCTTCC 203
Qy 390 GGTGGGCGATTTTGTGTCCTTATGTCGATTTGGGTTTGGGATGGGAGCATCGGTA 449
Db 204 AGCTATGATGATTCATCGCCCGCTGCTGTCATGCGCTGAGCGCGCGCGCATCGCA 263
Qy 450 ACTTATGATATGATGAACTTTGGGAGCGATGAACGATGTCCTTGGGCTTTTGTATTC 509
Db 264 ACTTATGATGATGCGAAGCTGTGGGAGCGGCAAGCGATGAGCCCTGGGCGCATGCTTC 323
Qy 510 CTAATGTTGGCCCATGCGCGCATCTTCAAGCTTTAATTCGCTTGAAGGCG 569

```

```

Db 324 CCAAGCGGCGCCGCTGCGCGGACATCCCTCGAAGTGTACCAAGTTGCGCCCTGAAGGCC 383
Qy 570 TGTCTCTGTTCTTATTCTTATTATGTTATTTATGTTAACTCTGCTCCGTAAGCAGCGTAT 629
Db 384 TGGCATCTTGTGATCTCTTGGGCGGCAACCCGATGCGAGGTGCAACGCGCTTGTCT 443
Qy 630 CCGAATGTTTATAGCTGATACGATACATTCCTGCTTCTTGTGAAATGCTCCGTGAGC 689
Db 444 CTGGCTGTGTGCTGCTGCTGCTACCGGATCTACCGCTTGTGATGGGGAATTCACCGTGTGC 503
Qy 690 CAGATGCTATTTGG--GTCGTGTTGTGCTTCATTTCAATGGGCGAATTCCTCTCT 746
Db 504 CGCATGCCACCTGCTTACTTCCATGATGATGCTGATGATGCGGCAAGAGCTGCGAG 563
Qy 747 TACCTATGATGATCATCGTATTT 770
Db 564 TACCGTGGAAATGCGCTTCT 587

RESULT 6
U83227 406 bp DNA linear GSS 21-FEB-2001
LOCUS U83227
DEFINITION U83227 Pseudomonas putida Mmb1: ATCC 23483 Pseudomonas putida
genomic clone transposon-tagged mutant UT502 similar to
prolipoprotein diacylglycerol transferase (Lgt), genomic survey
sequence.
ACCESSION U83227
VERSION U83227.1 GI:1786113
KEYWORDS GSS.
SOURCE Pseudomonas putida
ORGANISM Pseudomonas putida
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 406)
AUTHORS Caspi, R., Tebo, B.M. and Haygood, M.G.
c-type cytochromes and manganese oxidation in Pseudomonas putida
Mmb1
JOURNAL Appl. Environ. Microbiol. 64 (10), 3549-3555 (1998)
MEDLINE 98432765
PUBMED 9758766
COMMENT Contact: Caspi R
MBRD
Scripts Institution of Oceanography, UCSD
9500 Gilman Drive, La Jolla, CA 92093-0202, USA
Email: rcaspi@ucsd.edu
This sequence was obtained by automated sequencing from one strand
of the DNA. It was NOT confirmed by sequencing the second strand.
Therefore, some mistakes might be present.
Class: transposon-tagged.
FEATURES
source location/Qualifiers
1..406
/organism="Pseudomonas putida"
/mol_type="genomic DNA"
/strain="Mmb1; ATCC 23483"
/db_xref="taxon:303"
/clone="transposon-tagged mutant UT502"
/clone_lib="Pseudomonas putida Mmb1; ATCC 23483"
/note="manganese oxidizing bacterium"
ORIGIN
Query Match 4.0%; Score 115.6; DB 29; Length 406;
Best Local Similarity 59.8%; Pred. No. 3.8e-22;
Matches 234; Conservative 0; Mismatches 145; Indels 12; Gaps 2;
Qy 399 ATTTTGTGCCCCCTTATGTCATTCGTTGGGATGGGAGCATGCGTAACTTATGA 458
Db 15 ATTCTGTGCCCCCGCTGTGCGCATTCGCGCTGGGTCGCGCGGCGATGCGCACTTCATCA 74
Qy 459 ATAGGAATTTGGGAGCGATGAACGATGATGCTTGGGCTTTTGTATTTCC----- 510
Db 75 ACGCGAGCTGTGGGCAAAACCGACGATGCGCATGATGATCTTCCGCGCGTTCA 134
Qy 511 -TAATGTGGCCCACTGCGCGCATCTTCAAGCTTTATGAATTCGCTTGAAGGCG 569

```

Db 135 GGCACCCAGCCAGCTGCCACGCCCTGCACTGTAACGATTGCGCCCTGAGAGGCG 194
 QY 570 TGGTTCTGTTCTTATTCTTAATTGTTTATGTTAACTGCTCGGCTAGGAGGTTAT 629
 Db 195 TGGCACTGTTCCGATCTCTGCTGCTCTGCGCAAGCGCGCCGACCACTGCGGCTAT 254
 QY 630 CCGGACTGTTTATGCTGATAGGATACATTCGCTTCTTGGTAATAGTCCTGAGC 689
 Db 255 CCGGACTGTTGCGGCTGTTCTAGGCACTTCGCTTCACTGTCGAATTCGCGGCTGC 314
 QY 690 CAGATGCTCAGTTGGG---TCTGTTTGTGGCTTCAATTCATGGGCAATCTCTCT 746
 Db 315 CGGATCCAGCTGGGCTATATGCGCTGGGGCTGGCTGACCAATGGGTGAGATTCTGTGG 374
 QY 747 TACCTATGCTGATCATCGGTATTCTTATGAT 777
 Db 375 TGCCGATGATCTGCTGCGCTTGCGCTGAT 405

RESULT 7

LOCUS BZ561697 1034 bp DNA linear GSS 17-DEC-2002
 DEFINITION pac62-164_3389.y2 pac62-164 Pseudomonas aeruginosa genomic clone
 ACCESSION BZ561697
 VERSION BZ561697.1 GI:27182207
 KEYWORDS GSS.

SOURCE Pseudomonas aeruginosa
 ORGANISM Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.

1 (bases 1 to 1034)
 Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Haerlings, M.,
 Burns, V.L., Kaul, R. and Olsen, M.V.
 Whole-Genome-Sequence Variation among multiple Isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press
 Contact: Chris K. Raymond
 Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: ckraymond@u.washington.edu
 Classes: Shotgun.

JOURNAL

COMMENT

FEATURES

source Location/Qualifiers
 1..1034
 /organism="Pseudomonas aeruginosa"
 /mol_type="genomic DNA"
 /strain="2-164"
 /db_xref="taxon:287"
 /clone="pac62-164_3389"
 /clone_lib="pac62-164"
 /note="Clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 3.8%; Score 111.2; DB 28; Length 1034;
 Best Local Similarity 53.9%; Pred. No. 1.2e-20;
 Matches 342; Conservative 0; Mismatches 278; Indels 14; Gaps 5;
 QY 29 CTGCAATTTCCCAATATTGACCCCGATTGTTTTCATGAGCGGCTTACGCGGCGCTGG 88
 Db 370 CTGACCTATCCCGAGATCGATCGGTTGCGCTGCCATCGCCCGTGAAGATCCACTGG 429
 QY 89 TATGCTTATGATATTGTTGGGTTTCCTTTTGTATGTTGTTGGCAATGCGCCAGG 148
 Db 430 TAGGGCTATGATGCTGATCGGATCGCGGCGCTGCTGCTGCGCTGCGCCGAG 489
 QY 149 GATCGGCGGCGAGTGTGAGCGCTGAGCAAGTCTTGACTTGTATTTC---GCCGCG 205
 Db 490 AAGCGCTTGACCCGACCTGACCAAGAGCGCGCTTCCGACTGTGCTTGTGGTTGCG 549

QY 206 TTTTAAAGTGTATGATCGGTGCGAGTTGGTTATGTTATCTCTACAAATTTGATCTG 265
 Db 550 CTGGCGGCGGTATCTCCCGGTGGCGCCCTGGGCTAGTCTGTTCTAGAACCTGAGCGAG 609
 QY 266 TTCTTGTCTGACCTCTTTATTATTCAAAGTGTGACCTGGCGGCAATGCTTCCAGCGC 325
 Db 610 TACATGCGCAACCGACGCTGATCTTCAGAGCTGCGAAGGGCGGCAATGCTTTCAATGCG 669
 QY 326 GCGTATGGGTGTATACCGCCATGTTCTGTGATGCGCGTAAACCAAGCAAGCACTTC 385
 Db 670 GGGCTGCTCGGGTATGCTGCGGCTGCTGTTTTCGCAAGCGCCATGCGCAAGAGAC 729
 QY 386 TTGGTGTGGCGCA-----TTTGTGCCCCCTTAATGTCATTGCGTTTGGAGATGGAGC 440
 Db 730 TTCTTACGCTGAGGAGACTTCAATTCGCCCCCTGTGTCCTATGAGCCCTGGCGCGCGC 789
 QY 441 GTATCGTAACTTTATGAATAGTGAACCTTGGGAGAGAG--TAAAGATGTGCTTGG 497
 Db 790 GCATTGGGAAACTTTTATTAATGAACTTGTGGGCGAAGTCAATGATGAGCCCGCGGC 849
 QY 498 CTTTGTATTTCCCTTAATGTGCGCCCACTGCGCCGCAATCTTCAAGCTTATGAA-TTC 556
 Db 850 CAGGCGCTTTCCTCCCAACGTGCGCCCTTTCGCGGCGATCCCTTACATTTGCCAATTC 909
 QY 557 GCCTTAGAAGCGGTGCTCTGTTCTTTA--TTCTTAATGTTTATGTTAAGTAACTGTC 614
 Db 910 GCCTTGAAGCGGCGGCGCCGCTTTTATATCTTGTGGTTAGTTAACCCCAAGCCGCGC 969
 QY 615 CGCTAGGACGCTATCCGAGCTGTTTATGACTGG 648
 Db 970 CGCGCGCCCTTGCTTCCGCTCTTTCCGCGCTTG 1003

RESULT 8

LOCUS CD379073/c 878 bp mRNA linear EST 31-MAY-2003
 DEFINITION PTM004242 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum
 tricornutum cDNA 5', mRNA sequence.

ACCESSION CD379073
 VERSION CD379073.1 GI:31254687
 KEYWORDS EST.
 SOURCE Phaeodactylum tricornutum
 ORGANISM Phaeodactylum tricornutum

Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.

1 (bases 1 to 878)
 Scala, S., Carels, N., Falciatore, A., Chiusano, M.L. and Bowler, C.
 Genome properties of the diatom Phaeodactylum tricornutum
 Plant Physiol. 129 (3), 993-1002 (2002)

COMMENT
 MEDLINE
 PubMed
 Contact: Bowler C
 Laboratory of Molecular Plant Biology
 Stazione Zoologica 'Anton Dohrn'
 Villa Comunale, I-80121, Napoli, Italy
 Tel: 39 081 583 3268/3211
 Fax: 39 081 764 1355
 Email: chris@alpha.szn.it
 Diatom EST Database (http://aves.thegen.sznbowler.com)
 Seq primer: T3 backward
 POLVA=yes.

FEATURES
 source Location/Qualifiers
 1..878
 /organism="Phaeodactylum tricornutum"
 /mol_type="mRNA"
 /db_xref="taxon:2850"
 /cell_line="CCMP632"
 /clone_lib="Phaeodactylum tricornutum Uni-Zap XR"
 /note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2:
 Xho I"

ORIGIN

Query Match	3.8%;	Score 111;	DB 14;	Length 878;
Best Local Similarity	59.8%;	Pred. No. 1.2e-20;		
Matches 186;	Conservative 0;	Mismatches 125;	Indels 0;	Gaps 0;
QY	2474	ATCATTTTCTAAGGTAGGAGATCGACGACCGGAGAGAAATTGAAAAAATCGTGATGCTG	2533	
Db	332	ATGACTTCTACGAGAAAGATATGGCCACCGGACAAAGAAATTAACATATCGTGACGCTG	273	
QY	2534	GCACTGGCGGAAAGCAGCGCGGAACTCTTTCTTTACAGCGCATATGGCCCAAGGCTGACGAG	2593	
Db	272	GCACGAGGAAAGCAGCGCTCGAGCTGATGCCCATGCCCATTTGGCCCATAGCGAGATT	213	
QY	2594	GTATTGGTCACAGTGTAGTACCAATATTGGACCCCATCACATAGAAATGGCGGTTTCAAC	2653	
Db	212	GTTTGGTGACTGATGATGACCAATATTCCTCTCCGATGATCATGTAAATTCCTTGGCGGAC	153	
QY	2654	GGTAAACCCAGCGGCAACGAGACCAACAATTAATAGAGTACCGTGTCTTGAGATTAATC	2713	
Db	152	GAGACAGCGGATTCGACCAAGAGACCAACATTCGAGTGTGTATGAAGAAAGACTGAGC	93	
QY	2714	AGTGGCGTTGGCCACTAAACCAATCATCATCTGCAATTGGTGGGAAACAAATTCAAAT	2773	
Db	92	AATCAGTCGCAAGATATCAACCATATGACGACGAAATGGGATTCGTCTGCGCCGAAAC	33	
QY	2774	AGAACTTTGGC 2784		
Db	32	AACCTTCGCGC 22		
RESULT 9	CD382942/c			
LOCUS	CD382942	921 bp	mRNA	linear
DEFINITION	PTMM08112 Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum			
ACCESSION	CD382942			
VERSION	CD382942.1	GI:31258556		
KEYWORDS	EST.			
SOURCE	Phaeodactylum tricornutum			
ORGANISM	Phaeodactylum tricornutum			
REFERENCE	1 (bases 1 to 921)			
AUTHORS	Scala,S., Carels,N., Falciatore,A., Chisano,M.L. and Bowler,C.			
TITLE	Genome properties of the diatom Phaeodactylum tricornutum			
JOURNAL	Plant Physiol. 129 (3), 993-1002 (2002)			
MEDLINE	22111123.			
PUBMED	12114555			
COMMENT	Contact: Bowler C			
	Laboratory of Molecular Plant Biology			
	Stazione Zoologica 'Anton Dohrn'			
	Villa Comunale, I-80121, Napoli, Italy			
	Tel: 39 081 583 3268/3211			
	Fax: 39 081 764 1355			
	Email: chriss@alpha.azn.it			
	Diatom EST Database (http://aves.thagen.sznbowler.com)			
	Seq primer: T3 backward			
	POLYA=yes.			
FEATURES				
source	Location/Qualifiers			
	1..921			
	/organism="Phaeodactylum tricornutum"			
	/mol_type="mRNA"			
	/db_xref="taxon:2850"			
	/cell_line="CCMP632"			
	/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"			
	/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"			
ORIGIN				
Query Match	3.8%;	Score 111;	DB 14;	Length 921;
Best Local Similarity	59.8%;	Pred. No. 1.3e-20;		
Matches 186;	Conservative 0;	Mismatches 125;	Indels 0;	Gaps 0;
QY	2474	ATCATTTTCTAAGGTAGGAGATCGACGACCGGAGAGAAATTGAAAAAATCGTGATGCTG	2533	

Db	253	ATGACTTTTACAGGAAGAAGTATGGCCACCGACAGAAATTAAACATATCTGTGACGGTG	294
QY	2534	GCACCTGGCGAAAGACCGCGAAACTCTTCTTTACAGCGCATATGSCCAAGGCTGACGAGA	25933
Db	293	GCACACGAGGAAGAAGCACGCTCGACGTGATCCCATTCGCCATTTTGSCCATATGCGACGAT	234
QY	2594	GTATTTGGTTCACATGATATCCAAATATTTGGCACCCATCACATAGGAATTCGGCTTTCAAC	26533
Db	233	GTGTTGGTGCATGGAATGACCAATATTTGGCTCCCATGATCATGTAAATTTCTTGGCGAGC	174
QY	2654	GGTAAACCCACCGGGAACGAGACCAACAATATATAGAGTACACGTCGTCTGAGATTGAATC	2713
Db	173	GAGACAGCGGATCCGACCAAGAGACACAAATCGAGGTCGTTGTATGAGAAAGACTGGAGC	114
QY	2714	AGTGCCTGTGGCCATMAACCAATCATCATCTGCAATTTGGTGGGAAGCAAAATTCAAAT	2773
Db	113	AATACAGTCGGAAGATACCAACATTTAGACACCAATGGATTCGTGTGCGCGCAAC	54
QY	2774	AGAACTTTGGC	2784
Db	53	AACCTCCGCG	43
RESULT 10			
CD376447/c		855 bp	mRNA linear EST 31-MAY-2003
LOCUS	PTM001614	Phaeodactylum tricornutum Uni-Zap XR Phaeodactylum	
DEFINITION	tricornutum cDNA 5', mRNA sequence.		
ACCESSION	CD376447		
VERSION	CD376447.1	GI:31252061	
KEYWORDS	EST.		
SOURCE	Phaeodactylum tricornutum		
ORGANISM	Phaeodactylum tricornutum		
REFERENCE	Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae; Bacillariophycidae; Naviculales; Phaeodactylaceae; Phaeodactylum.		
AUTHORS	1 (bases 1 to 855)		
TITLE	Scala,S., Carels,N., Falcitore,A., Chiusano,M.L. and Bowler,C.		
JOURNAL	Genome properties of the diatom Phaeodactylum tricornutum		
MEDLINE	Plant Physiol. 129 (3), 993-1002 (2002)		
PUBMED	22111123		
COMMENT	12114555		
	Contact: Bowler C		
	Laboratory of Molecular Plant Biology		
	Stazione Zoologica 'Anton Dohrn'		
	Villa Comunale, I-80121, Napoli, Italy		
	Tel: 39 081 583 3268/3211		
	Fax: 39 081 764 1355		
	Email: chris@alpha.szn.it		
	Diatom EST Database(http://aves.thagen.szbowler.com)		
	Seq primer: T3 backward		
	POLYA=Yes.		
FEATURES			
source	Location/Qualifiers		
	1..855		
	/organism="Phaeodactylum tricornutum"		
	/mol_type="mRNA"		
	/db_xref="taxon:2850"		
	/cell_line="CCMP632"		
	/clone_lib="Phaeodactylum tricornutum Uni-Zap XR"		
	/note="Vector: Uni-Zap XR vector; Site_1: Eco RI; Site_2: Xho I"		
ORIGIN			
Query Match	3.7%;	Score 106.2;	DB 14; Length 855;
Best local Similarity	58.8%;	Pred. No. 3.1e-19;	
Matches 183;	Conservative 0;	Mismatches 128;	Indels 0; Gaps 0;
QY	2474	ATCATTTCTTAGAGGATGAGATGACGACCGCGAGAGATTGAAGAAAATCGTGATGGTG	25333
Db	323	ATGACTTTTACAGGAAGAAGTATGGCCACCGACAGAAATTAAACATATCTGTGACGGTG	264
QY	2534	GCACCTGGCGAAAGACCGCGAAACTCTTCTTTACAGCGCATATGSCCAAGGCTGACGAGA	25933

Db 263 GCACGACGGAAGACGCTGAGTGGATCCCATGCCCCATTTGGCCCATAGCAGCATG 204
 QY 2594 GTATTGTCACAGTATGATCAATATTTGGACACCATCATCAGATTCGGTTTCAACC 2653
 Db 203 GTGTGGTGCATGAAATGATACCAATATTCGTCCTCCATGATCATGTAATATCTTGGCGGAGC 144
 QY 2654 GGTAAACCAACCGGCAACGAGCAACCAATATAGAAATGACCGTGTGGATTTGAATC 2713
 Db 143 GAGACAGCGGATCCGACCAAGACCAACATCGAGGTAGTGTGTAAGACGACGTAAGC 84
 QY 2714 AGTCCGCTTGCACATTAACCATCATCATCTCGCAATTTGGTGGAGCAATTTCAAT 2773
 Db 83 AATAACATCGCAAGATATCAACCAATTAGACGAGCATGAGTATGCTGTGTCGCGCAAC 24
 QY 2774 AGACCTTGGC 2784
 Db 23 AACTCTCCGAC 13

RESULT 11
 CAS88286/c 843 bp mRNA linear EST 19-NOV-2002
 LOCUS hab9d11.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone
 IMAGE:6351812 5' similar to TR:Q91237 Q91237 NA/PI COTRANSPORT
 SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION
 VERSION CAS88286
 KEYWORDS EST
 CAS88286.1 GI:25132864

SOURCE
 ORGANISM Takifugu rubripes (Fugu rubripes)
 Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphi; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Takifugu.
 1 (bases 1 to 843)
 Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M.,
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
 Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 WashU Zebrafish EST Project 1998
 Unpublished (1998)
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.edu

TITLE
 JOURNAL
 COMMENT

FEATURES
 source

1. 843
 /organism="Takifugu rubripes"
 /mol_type="mRNA"
 /db_xref="caxon:31033"
 /clone="IMAGE:6351812"
 /cissue_type="gut"
 /dev_stage="adult"
 /lab_host="PH108 (T1 phage-resistant)"
 /clone_lib="Fugu UT6 adult gut"
 /note="Vector: pBluescript-Pu; Site 1: pLIM I
 (CCANNNNNTGG); Site 2: pLIM I (CCANNNNNTGG); BamHI-SmaI
 sites were converted to BamHI-pLIM1-SfiI-pLIM1 sites (SmaI
 is destroyed). Other part of the vector is untouched. The

ORIGIN
 Query Match 3 6%; Score 104; DB 14; Length 843;
 Best Local Similarity 58.9%; Pred. No. 1.4e-18;
 Matches 198; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

2427 GCAGAGCGAAACGACCGCATGTGAGACTTTCTTAATAATGCGCAACATCATTTCTAGAG 2486
 Db 721 GGAAGAGCTAATCAAGAACTGTGGAACCTTGAAGAATCTCTGTAGCACCTCCAGG 662
 QY 2487 GTAGAAAGATCGACCGCGAGAGATTGAAGAAATCGTGAATGCTGTGCACTGGCGAAG 2546
 Db 661 GCAGAAAGATCAACACCGACGCGCATTAAGAAATGCTGTGAGCGGTCCCGCAGCAAGG 602
 QY 2547 CACGCGGAACTCTTTTACGCGCATATGCGCAAGCTGACGAGATTTGTGACAG 2606
 Db 601 CCTGCGGAATCTGTTTGATCCCGCGCTGATCATGCGCCAGATGTTGTGACAG 542
 QY 2607 TAGTACCAATATTTGACCCCATCATCATAGATGCGGTTT---CAACCGTAACCCAC 2663
 Db 541 AGTCCCGAATGTGGCGCCCATGATATCGGACCGCGGTCTGACATCATGATTTCCAG 482
 QY 2664 CGGCAAGACGACCAACATATATAGAAATGACCGTCTTGAGATTTGAATGATGCGGTTG 2723
 Db 481 AGGACACCATGTGACCAACATATAGAGAGAAAGTCTGAGAGCTTGACACGACGATGA 422
 QY 2724 CCACATAACCATCATCATCTGCAATTTGGTGGG 2759
 Db 421 CTACACCCCGATCCAGCCGACGACGACGAGGTTG 386

RESULT 12
 CAS88344/c 813 bp mRNA linear EST 19-NOV-2002
 LOCUS hab50c01.y1 Fugu UT6 adult gut Takifugu rubripes cDNA clone
 IMAGE:6351745 5' similar to TR:Q91237 Q91237 NA/PI COTRANSPORT
 SYSTEM PROTEIN. ; mRNA sequence.

ACCESSION
 VERSION CAS88344
 KEYWORDS EST
 CAS88344.1 GI:25132922

SOURCE
 ORGANISM Takifugu rubripes (Fugu rubripes)
 Takifugu rubripes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphi; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Takifugu.
 1 (bases 1 to 813)
 Clark, M., Johnson, S. L., Lehrach, H., Lee, R., Li, F., Marra, M.,
 Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
 Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
 Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 WashU Zebrafish EST Project 1998
 Unpublished (1998)
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrafish@wustl.edu

TITLE
 JOURNAL
 COMMENT

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library
 constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi
 and S. Watabe (University of Tokyo, Institute of Medical Science
 and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA
 Sequencing by: Washington University Genome Sequencing Center Clone

distribution: Fugu clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov

Seq primer: T3 RT from Amersham
High quality sequence stop: 497.
Location/Qualifiers

FEATURES

source

```
1. 813
/organism="Takifugu rubripes"
/mol_type="mRNA"
/db_xref="taxon:31033"
/clone="IMAGE:6351745"
/rissue_type="gut"
/dev_stage="adult"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="Fugu UT6 adult gut"
/notes="Vector: pBluescript-PL; Site 1: PflM I (CCANNNNNTGG); Site 2: PflM I (CCANNNNNTGG); BamHI-SmaI sites were converted to BamHI-PflM-SfiI-PflM sites (SmaI is destroyed). Other part of the vector is untouched. The cDNA is inserted between two PflM sites in T3 (5') to T7 (3') direction. Library materials provided by G. Elgar (UK MRC HGP-RC) and constructed and donated by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Watabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology)."
```

ORIGIN

```
Query Match 3.6%; Score 103.8; DB 14; Length 813;
Best Local Similarity 59.2%; Pred. No. 1.6e-18;
Matches 196; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

OY 2432 GGCAGAAAGAGCCAGTGTGAGACTTCTCTAAATGCAAAACATCATTTCTAGAGTGG 2491
DB 713 GACGTAATCAGAAATGGGTGAACCTGTAAAGAACTCTGTAGCCACCTCCAGAGCAGA 654
OY 2492 AAGATCAGCAGCCGAGAAATGTAATAATCGTGAATGTGGCACTGGCCAGAACACAGG 2551
DB 653 AGATCAACACCGACAGCAAGTTAAAGAGTGTGAGCGTGGCCCCAGCAAAAGGCCCTG 594
OY 2552 CGAAACTCTTCTTACAGCGCATATGGCCAGAGTGAAGAGTATTTGTCACTACTATA 2611
DB 593 CGGAACCTGTTTCATCCCGCCCGCTGATCATGCGCAGATGTGTGTGTGACAGAGTTC 534
OY 2612 CCAATATTGGACCCATCAGCATAGGAATCGCGGTTT---CAACGGTAAACCCAGCGCA 2668
DB 533 CCGATGTGGCGCCCATATATATGCGCACCGGGCTGTGACATCCAGTAATTCAGAGAAC 474
OY 2669 ACGAGACCAACAAATATAGAAAGTCAACGTCCTTGAAGATTGAATCAATGCGGTGCACT 2728
DB 473 ACCATGTGACCAACATATAGAGAGAGAGTGTGAGCTGTGACGAGACGCGTACTAAC 414
OY 2729 AAACCAATCATCATCTGCAATTGGGTGGG 2759
DB 413 ACCCGATCACCGCCAGCCACAGCGGTTGG 383
```

RESULT 13

```
CA475083/c CA475083 828 bp mRNA linear EST 12-NOV-2002
LOCUS AGENCOURT 10670196 NCI CGAP ZKId1 Danio rerio cDNA clone
DEFINITION IMAGE:6797422 5', mRNA sequence.
ACCESSION CA475083
VERSION CA475083.1 GI:24931435
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 828)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: LLM14310 row: k column: 21
High quality sequence stop: 433.
Location/Qualifiers

FEATURES

source

```
1. 828
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6797422"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI CGAP ZKId1"
/notes="Organ: Kidney; Vector: PCMV-SPOrt6.1; Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Constructed by J. Wang (Research Genetics, Invitrogen Corp) from tissue donated by L. Zon (Harvard University). Note: this is a NCI CGAP library."
```

ORIGIN

```
Query Match 3.5%; Score 103.2; DB 14; Length 828;
Best Local Similarity 60.5%; Pred. No. 2.4e-18;
Matches 187; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

OY 2465 ATGCCAAACATCATTTCTAGAGTGAAGATGACACCGGAGAAATGAAAAATCG 2524
DB 521 ATGCTGTGATGTTTCCAGGCGAGACGACGACACTGACACAGTTAAACAAATCG 462
OY 2525 TGGATGTGSCACTGGCGAAAGCAACGGCGAAATCTTCTTACAGCGCATATGGCCAAAG 2584
DB 461 TGCACTGTGGACACGACAAAGCTCGTNTGAATCTCTCTCCCTGCTGCATGAGA 402
OY 2585 CTGACGAGATATTTGTTCAGATGTACCAATATTTGCAACCATCACCATTAGAAATGGCG 2644
DB 401 GCCACGATGTGTTGTGACAGATGTGCCAATGTTTGAAGCCATGATGAGAGAGCG 342
OY 2645 GTT---TCAACCGGTTAAACCCAGCGCAACGAGCAACAAATATATAGAACCGTGCTT 2701
DB 341 GATCCACATCTACACAAATCCAGAGAAACGACGTGACAAATAGAGGTAGAGTGTG 282
OY 2702 GAGGATTGAATCAGTCCGTTGCCACTTAACCAATCAATCAATCTGCAATTTGGTGGAA 2761
DB 281 GAACCTTGACAGAGACAGTAACAGATATTCACACACAGTCCCGCACAGGGTTAGAA 222
OY 2762 GCAATTCGA 2770
DB 221 AGAACCGCA 213
```

RESULT 14

```
BM039803/c BM039803 685 bp mRNA linear EST 19-OCT-2002
LOCUS BM039803 Nori Satoh unpublished cDNA library, blood cells Ciona
DEFINITION intestinalis cDNA clone cibd041k19 5', mRNA sequence.
ACCESSION BM039803
VERSION BM039803.1 GI:24140498
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 685)
AUTHORS Satou, Y., Satake, M., Azumi, K., Nonaka, M., Shin-i, T., Kohara, Y. and Satoh, N.
TITLE Expressed genes in Ciona intestinalis (2002)
```


JOURNAL
COMMENT

Unpublished (2002)
Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source

Location/Qualifiers
1..685
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cibd041k19"
/tissue_type="blood cells"
/clone_lib="Nori Satoh unpublished cDNA library, blood cells"

ORIGIN

Query Match 3.3%; Score 96; DB 13; Length 685;
Best Local Similarity 56.5%; Pred. No. 2.8e-16;
Matches 199; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 2474 ATCATTTCTAGAGTGAAGATGACAGCCGCGAGAAAGATTGAAAAATCGTGATGTG 2533
DB 616 ATTTGTTCCAGTGAAGTAAATGAGCACTACTGCAACGTGAAAGAAATCGTGACAGTC 557

QY 2534 GCACTGGCGAAAGCAGCGGAAACTCTTTTACAGCCGATAGGCCAAGCTGACGAGA 2593
DB 556 GCTCCCTCGGTAGCCTTTCTGAATTCGTTTGTGCGGATATGCGCAAGATACGATG 497

QY 2594 GTATTGTCACAGTAGTACCAATATTGGACCCATCAGCATAGGAATGCGGTTTCAACC 2653
DB 496 GTGTTTGATGGAATTCGAATGTTGGCCCATGACATGGCGATGCGGTTTACA 437

QY 2654 GGTAAACCCACCG---GCAACGAGACCAACAATATAGAGTCAACCTGCTTGAGATTGA 2710
DB 436 GTCAATGCGCGCTCACCACCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 377

QY 2711 ATCAGTCCGTTGCCATTAACCAATCATCATCTCTGCAATTGGGTGGAGCAAAATTGA 2770
DB 376 AAGATGAGAGTGAAGATTCCTATGACAGAGCCCGCATGATGATTATCAGGAGAGGAG 317

QY 2771 AATGAACTTTGGCTTGATGCGCGGTGCGCATTTAAACCGCTGCGGACCA 2822
DB 316 AAAAAAGTGTGTGATTTGGCGGTGAGGACTTTGAAGAGTTGCCAAGCA 265

RESULT 15
BM039194/c

LOCUS BM039194 686 bp mRNA linear EST 19-OCT-2002
DEFINITION BM039194 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA cibd039004 5', mRNA sequence.

ACCESSION
BM039194

VERSION BM039194.1 GI:24139889

KEYWORDS

EST

SOURCE

ORGANISM

Ciona intestinalis

Ciona intestinalis

Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.

Sato, Y., Satake, M., Azumi, K., Nonaka, M., Shin-1, T., Kohara, Y. and

Sato, N.

Expressed genes in Ciona intestinalis (2002)

Unpublished (2002)

Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoh@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers

source

1..686
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cibd039004"
/tissue_type="blood cells"
/clone_lib="Nori Satoh unpublished cDNA library, blood cells"

ORIGIN

Query Match 3.3%; Score 96; DB 13; Length 686;
Best Local Similarity 56.5%; Pred. No. 2.8e-16;
Matches 199; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 2474 ATCATTTCTAGAGTGAAGATGACAGCCGCGAGAAAGATTGAAAAATCGTGATGTG 2533
DB 615 ATTTGTTCCAGTGAAGTAAATGAGCACTACTGCAACGTGAAAGAAATCGTGACAGTC 556

QY 2534 GCACTGGCGAAAGCAGCGGAAACTCTTTTACAGCCGATAGGCCAAGCTGACGAGA 2593
DB 555 GCTCCCTCGGTAGCCTTTCTGAATTCGTTTGTGCGGATATGCGCAAGATACGATG 496

QY 2594 GTATTGTCACAGTAGTACCAATATTGGACCCATCAGCATAGGAATGCGGTTTCAACC 2653
DB 495 GTGTTTGATGGAAGTTCGAATGTTGGCGCCCATGACATGAGGATGCGGTTTACA 436

QY 2654 GGTAAACCCACCG---GCAACGAGACCAACAATATAGAGTCAACCTGCTTGAGATTGA 2710
DB 435 GTCAATGCGCGCTCACCACCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 376

QY 2711 ATCAGTCCGTTGCCATTAACCAATCATCATCTCTGCAATTGGGTGGAGCAAAATTGA 2770
DB 375 AAGATGAGAGTGAAGATTCCTATGACAGAGCCCGCATGATGATTATCAGGAGAGGAG 316

QY 2771 AATGAACTTTGGCTTGATGCGCGGTGCGCATTTAAACCGCTGCGGACCA 2822
DB 315 AAAAAAGTGTGTGATTTGGCGGTGAGGACTTTGAAGAGTTGCCAAGCA 264

Search completed: July 31, 2004, 16:59:42
Job time : 4830 secs

This Page Blank